(9) FINAL REPORT. LEVEL II

AD A073449

(6) METHODS FOR FACTOR SCREENING IN COMPUTER SIMULATION EXPERIMENTS.

20000731231

By (10)
Douglas C. Montgomery

Prepared for (15)
The Office of Naval Research
Contract N00014-78-C-0312

(11) March 1979 (12) 90p.

GEORGIA INSTITUTE OF TECHNOLOGY
SCHOOL OF INDUSTRIAL & SYSTEMS ENGINEERING
ATLANTA, GEORGIA 30332

409 390

METHODS FOR FACTOR SCREENING IN COMPUTER SIMULATION EXPERIMENTS

Douglas C. Montgomery
School of Industrial and Systems Engineering
Georgia Institute of Technology

Prepared for the Office of Naval Research
Under Contract N00014-78-C-0312

March, 1979

SUMMARY

The use of a computer simulation model may be viewed as an experiment in which a set of k input variables are combined to produce at least one output or response variable. As in any experimental situation, the design of a computer simulation experiment is important. In general, not all k input variables or factors will be equally important in their effect on the response variable(s). It is very common to find that only a subset, say $g < k$, of the original k factors are important in explaining the response. We usually do not know the value of g, or which g factors are important.

The problem of experimentation and analysis to discover the size and composition of the subset of active factors g is called the factor screening problem. It is important to accurately identify the set of active factors. Failure to identify an active factor can result in serious bias in the analysis and conclusions drawn from the model, if that factor is subsequently ignored. Conversely, experimentation with negligible factors is undesirable as it consumes the resources of experimentation needlessly.

This report contains a survey of the available statistical methodology useful in factor screening. It also discusses the relative merits of each approach, and provides guidelines for the development of a factor screening strategy. Several examples are presented that demonstrate the construction of factor screening experiments, and the interpretation of the results of such experiments.

Three types of factor screening situations may be identified. The first case is the designed experiment situation; that is, a situation in which an experiment is designed and conducted with the primary objective of discovering

the set of active factors. The use of designed experiments in factor screening is particularly important, as designed experiments allow assessment of main effects and interactions independent of other effects that may be present in the mode. Designed experiments also often allow the incorporation of variance reduction methods. Finally, they usually admit a relatively simple statistical analysis.

The major classes of factor screening designs discussed in this report include:

1. The 2^{k-p}_{III} and 2^{k-p}_{IV} fractional factorial designs
2. Supersaturated designs
3. Group screening designs
4. Irregular fractional factorials

A logical screening strategy involving those designs is developed. The selection between designs is based on consideration of the extent of aliasing of interactions and the severity of assumptions required to produce a unique analysis of the data. In particular, it is shown that group screening followed by the use of a 2^{k-p} fractional factorial design is often an optimal screening approach. Variance reduction methods for these designs are discussed, based on common and antithetic random number streams. Other problems discussed include the composition of the groups in group screening and selecting levels for negligible factors in subsequent experiments.

A second major type of screening study is the undesigned case. These situations occur when there are data available from previous simulation experiments with the model, and decisions regarding the identification of active factors must be made using these data. It is unlikely that these runs will conform to any standard factor screening design. However, in these cases, the method of least squares can be used to fit an appropriate regression model to the data, and factor screening decisions can often be made using this model.

The usual nonorthogonality of such undesigned data makes the interpretation of these models difficult. Standardized regression coefficients can be used to simplify the interpretation, although this still does not solve the problems created by a nonorthogonal data set. Several measures of nonorthogonality are introduced, including variance inflation factors and conditioning numbers, and the use of these measures in assessing the problems in interpreting individual regression coefficients is discussed. In cases of extreme nonorthogonality, parameter estimation methods other than least squares are recommended.

The third type of factor screening study involves augmenting an available data set with a small number of new runs. The question of where these additional runs should be conducted is discussed. Two design augmentation methods are proposed, one based on minimizing the variance of the parameter estimates, and the other designed to minimize the bias resulting from factors thought to be negligible.

This work was supported by the Office of Naval Research (ONR) under contract N00014-78-C-0312. I am grateful to Dr. Thomas C. Varley of ONR for his advice and encouragement.

CONTENTS

1. INTRODUCTION

1-1. Uses of Simulation

Many problems in operations research are too complex to be modeled and analyzed entirely by mathematical methods. Computer simulation is widely used in the study of such problems. Typical problem areas in which computer simulation has been successfully employed include queueing, inventory, scheduling, quality control/reliability analysis, and maintenance and repair activities. The military has made extensive use of computer simulation to analyze complex combat processes, as well as supply and logistics activities.

A computer simulation may be viewed as an experiment in which a set of controllable input or independent variables are combined to produce at least one output variable, usually called the dependent variable or response. In performing a computer simulation experiment, the analyst will usually have one of two objectives in mind:

1. Investigate the relationships between the independent variables and the response, determining, if possible, which factors exert the greatest effect on the response, and the extent of interaction between the factors.

2. Determine the set of factor levels that, over some appropriate region of interest, optimize the response(s).

As in an experiment, the design of a computer simulation experiment is an important aspect of the investigation. The use of formal experimental design methods in computer simulation results in significant advantages to the analyst, including simplicity of data interpretation and (usually) economic efficiency with respect to the total number of simulation runs required. For

background reading in experimental design, consult Cochran and Cox [1957], Davies [1956], Hicks [1973], Montgomery [1976], or John [1971]. For discussion of the specifics of applying experimental design methodology to computer simulation, see Burdick and Naylor [1966], Fishman [1973], Hunter and Naylor [1970], Ignall [1972], Kleijnen [1975a, part II], [1977], and Montgomery and Evans [1975].

1-2. The Need for Factor Screening

We shall assume that a computer simulation model may be described by a set of k controllable input variables or factors. These factors are generally of two types:

1. Factors that are controllable or subject to design in the "real world" system being modeled, such as inventory reorder quantities, service rates, or the rate of fire of a weapons system.

2. Factors that are not controllable in the real system, such as demand, weather effects, or the location of enemy troops or equipment. For purposes of conducting the experiment, however, all k factors will be assumed to be controllable in the simulation; that is, we may induce desired weather effects, or control the movements of an enemy submarine.

In general, not all of these k factors will be equally important with respect to their effect on the response variable(s). The factors may range in importance from highly important to negligible. It is very common to find that only a subset, say $g < k$, of the original k factors are important in explaining the response variable. However, generally, we do not know the value of g, nor do we know which g factors are important. This situation is discussed by Jacoby and Harrison [1962], who state that the problem is frequently encountered in computer simulation.

The problem of experimentation to discover the size and composition

of the subset of active factors is called the factor screening problem. It is important that the set of active factors be accurately determined. Failure to identify an active factor can lead to serious bias in the analysis and conclusions drawn from a model, if that factor is ignored in subsequent experiments. On the other hand, experimentation with negligible factors is undesirable as it consumes the resources of experimentation needlessly, and may increase the noise level in the data to the point when real effects are more difficult to discover. For example, many of the optimization techniques applied to computer simulation models decrease rapidly in efficiency as the number of independent variables increases. Clearly, identification of the set of active factors plays a critical role in the successful use of this methodology.

Factor screening methods can be profitably employed at two places during the development and use of a computer simulation model. They can be employed at the model design and development stage. Applied at this stage, screening methods can affect the choice of variables used in the model and hopefully simplify the architecture of the final model. This may require experimentation with components or subroutines of the model, or, when practical, experimentation with the real-world system. When used in this manner, factor screening could contribute significantly to reducing the running time of a simulation model, if negligible factors can be identified. Factor screening is also applicable to a complete simulation model, although it is unlikely that any major simplification of the model structure will result. However, the total number of computer runs that are to be made in exercising the model may be substantially reduced if some factors are not active.

This report contains a summary of the available statistical methodology useful in factor screening. It also discusses the relative merits of each

approach, and provides guidelines for development of a screening strategy. Other questions, including the implementation of variance reduction methods, choice of levels for factors thought to be negligible, and some details of parameter estimation in linear statistical models are also discussed.

1-3. Factors, Levels, and Parameter Estimation

Suppose that $x_1, x_2, \ldots, x_k$ are the controllable factors in a computer simulation experiment and y is the (single) response. We assume that the general structure of the simulation is such that it can be expressed in the form

$$y = f(x_1, x_2, \ldots, x_k) + \varepsilon. \tag{1-1}$$

In this equation, f is a functional relationship that determines the mean value

of the response y, and ε is an error term such that $E(\varepsilon) = 0$. In factor screening problems it is almost always sufficient to assume that f is linear in the unknown parameters that relate the response to the factors. For example, one possible model would be

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon \qquad (1\text{-}2)$$

where $\beta_0, \beta_1, \ldots, \beta_k$ are unknown parameters.

To perform an experiment with this system, we must choose a set of values or levels for each factor, and then run the computer simulation model at some subset (or possibly the full set) of the factor level combinations. The choice of the number of levels of each factor and their spacing when the factor is continuous (or approximately so) is important. Generally, we should be guided by the information we have about the likely effect of that factor on the response y.

In most factor screening experiments, we are simply attempting to determine the effect of the factor, not necessarily trying to develop a useful predictive or interpolative equation. Consequently, a relatively small number of factor levels is generally employed. Often two levels, arbitrarily called high and low, are sufficient. For example, in Figure 1 we have illustrated the behavior of y as a function of the factor x. Although y and x are related in a complex nonlinear manner, the use of two levels for x will be sufficient to measure the effect of x. However, in cases where extreme curvature is present in the functional relationship, more than two levels will be necessary. Rarely, however, would more than three or four levels of the factor be employed in a factor screening study. The need for more than a small number of levels often indicates that the region of exploration for x is too large.

The spacing of factor levels is also important. Levels should be far enough apart to measure anticipated effects, but not so far as to cause nonlinearities in the functional relationship to distort or mask significant effects. For example, consider Figure 2. If the low and high levels of x are x_1 and x_2, respectively, then (depending on the amount of noise) it is highly unlikely that the effect of x on y will be discovered. On the other hand, if the low and high levels are x_1 and x_4, then the curvature in the functional relationship will likely mask the true effect of x. The choice of x_1 and x_3 (or x_3 and x_4) as low and high levels of x will reveal that x has a significant effect on y. Neither case, however, would be sufficient for defining the effect of x so that a predictive or interpolative equation valid over the entire range $x_1 \leq x \leq x_4$ could be developed.

The effect of a factor may be defined as the change in response y produced by a change in the levels of the factor. This is usually called a main effect. For example, consider the data in Table 1, which presents information obtained from an experiment with two factors x_1 and x_2. The main effect of x_1 is the difference between the average response at the high level of x_1 and the average response at the low level of x_1, say

$$\frac{50+20}{2} - \frac{42+10}{2} = 9$$

That is, the average response increase upon changing from the low to the high level of x_1 is 9 units. Similarly, the main effect of x_2 is

$$\frac{50+42}{2} - \frac{20+10}{2} = 31.$$

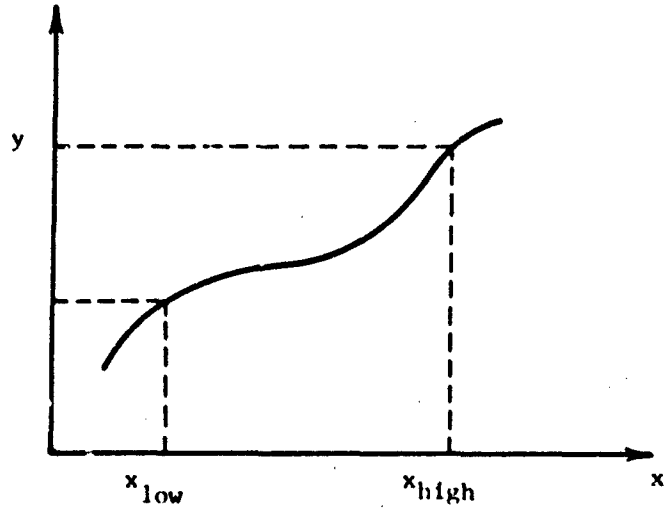

Figure 1. Use of Two Factor Levels to Model the Effect of x on y

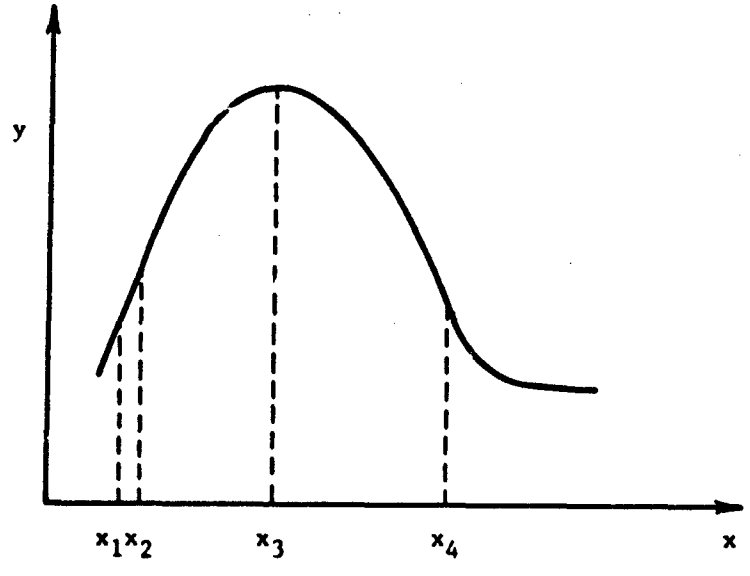

Figure 2. The Spacing of Factor Levels

Table 1

Data For a Factorial Experiment

		x_2 low	x_2 high
x_1	low	10	42
	high	20	50

The experimental design in Table 1 is a factorial design; that is, a design in which all possible factor level combinations are run. Furthermore, there is only one observation in each cell (we say the design is replicated once). Most screening designs are factorial designs.

Now consider the data in Table 2. Here the effect of x_1 is

$$\frac{30+20}{2} - \frac{42+10}{2} = -1$$

which implies that the x_1 effect is small. However, inspection of Table 2 reveals that the x_1 effect is not negligible, it just depends on the level of factor x_2. For example, at low x_2 the x_1 effect is

$$20 - 10 = 10$$

and at high x_2 the x_1 effect is

$$30 - 42 = -12.$$

Table 2

A Factorial Experiment

x_1 \ x_2	low	high
low	10	42
high	20	30

This is an example of an interaction between two factors. More specifically, it is a two-factor interaction. Most screening studies have to make certain assumptions about the types of interactions that are likely to be present in the system in order to design an economically efficient experiment. In general, factor screening attempts to sort out the main effects and low-order interactions that drive the system.

The method of least squares can be used to estimate the main effects and interactions. Suppose that we can describe the system by a linear statistical model, say

$$y_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \varepsilon_i, \quad i=1,2,\ldots,n \tag{1-3}$$

where y_i is the i^{th} response, x_{ij} is the i^{th} level of factor j, and β_j, $j=1,\ldots,k$ are unknown parameters. Letting $\underline{y} = (y_1,y_2,\ldots,y_n)'$, $\underline{\beta} = (\beta_0,\beta_1,\ldots,\beta_k)'$, $\underline{\varepsilon} = (\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)'$, where the prime denotes transpose, and letting X denote an $n \times (k+1)$ matrix whose first column is all ones and whose $(i,j+1)^{st}$ element is x_{ij}, then it is well-known that (1-3) can be written as

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon} \tag{1-4}$$

The least squares estimators of $\underline{\beta}$ are given by the solution to the normal equations

$$(X'X)\hat{\underline{\beta}} = X'\underline{y}, \tag{1-5}$$

or

$$\hat{\underline{\beta}} = (X'X)^{-1}X'\underline{y} \tag{1-6}$$

assuming that $(X'X)^{-1}$ exists.

To illustrate, consider the data in Table 1, and assume that the high and low levels of x_1 and x_2 can be represented by +1 and -1, respectively. Then (1-3) becomes

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i, \quad i=1,2,3,4.$$

We have assumed that x_1 and x_2 do not interact. Then, in matrix notation, we have for (1-4),

$$\begin{matrix} & & x_1 & x_2 \end{matrix}$$

$$\begin{bmatrix} 10 \\ 20 \\ 42 \\ 50 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}$$

The normal equations are

$$(X'X)\hat{\underline{\beta}} = X'\underline{y}$$

$$4I_3 \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 122 \\ 18 \\ 62 \end{bmatrix},$$

and the least squares estimates of the parameters in the model are

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 30.50 \\ 4.50 \\ 15.50 \end{bmatrix}$$

Note that the least squares estimates of the parameters are exactly half the main effects of x_1 and x_2; that is,

$$\hat{\beta}_1 = 4.50$$

$$\hat{\beta}_2 = 15.50$$

The parameter $\hat{\beta}_0 = 30.50$ is called the grand mean.

If we wished to incorporate interaction into this analysis, we would define the model as

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{i2} x_{i1} x_{i2} + \varepsilon_i, \quad i=1,2,3,4.$$

It is readily verified that

$$\begin{bmatrix} 10 \\ 20 \\ 42 \\ 50 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_{12} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}$$

and the normal equations become

$$4 I_4 \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_{12} \end{bmatrix} = \begin{bmatrix} 122 \\ 18 \\ 62 \\ -2 \end{bmatrix}$$

The least squares estimates of the parameters become

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_{12} \end{bmatrix} = \begin{bmatrix} 30.50 \\ 4.50 \\ 15.50 \\ -0.50 \end{bmatrix}$$

From examining the estimates of the effects, we conclude that both factors exert large (positive) main effects, while the two-factor interaction between those factors is negligible.

Users of statistically designed experiments are accustomed to analyzing the resulting data by relatively formal methods, such as the analysis of variance. In factor screening problems this is usually not done and the least squares estimates of the model parameters (or the effects) usually allow

significant factors to be identified. Often it is not practical to conduct a formal analysis of variance because of the small number of degrees of freedom that remain for error.

1-4. Designed and Undesigned Screening Experiments

The objective of a factor screening study is to discover as much as possible about the factors that significantly affect the response. Designed experiments are particularly useful in factor screening, as they allow assessment of effects and interactions independent of other effects present in the model, they often allow the incorporation of variance reduction methods, and they usually admit a relatively simple statistical analysis. However, screening is still possible in the undesigned case such as where there is data available from previous simulation runs. Once again, the method of least squares is useful here, although the usual nonorthogonality of such undesigned data makes the interpretation problem somewhat more difficult. Section 2 of this report will deal with designed screening studies, and Section 3 will discuss some aspects of undesigned screening situations, including the intermediate case in which some observations can be added to an existing data set.

In both cases, the method of least squares will be used for parameter estimation. We now state some useful results concerning least squares analysis of the general linear model. The model is

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon} ,$$

where y is $(n \times 1)$, x is $(n \times p)$, $\underline{\beta}$ is $(p \times 1)$, and $\underline{\varepsilon}$ is $(n \times 1)$. Note that the number of observations n must at least equal the number of parameters p. The least squares estimator of $\underline{\beta}$ is

$$\underline{\beta} = (X'X)^{-1}X'\underline{y} \ . \qquad (1\text{-}7)$$

If $E(\underline{\varepsilon}) = \underline{0}$ and the model is correct then the least squares estimators are unbiased; that is

$$E(\hat{\underline{\beta}}) = \underline{\beta} \ .$$

If the errors are uncorrelated with constant variance σ^2 then the covariance matrix of the least squares estimator is

$$\text{Cov}(\hat{\underline{\beta}}) = \sigma^2(X'X)^{-1} \qquad (1\text{-}8)$$

Note that the assumption of independent observations with constant variance will likely not hold in a simulation experiment. In fact, there are cases where the choice of variance reduction strategy induces a correlative structure between the observations. In cases where the assumption of uncorrelated errors with constant variance does not hold, the method of <u>weighted least squares</u> is useful. If V is a matrix of weights (chosen proportional to the variances and covariances of the errors) then the weighted least squares estimator of $\underline{\beta}$ is

$$\hat{\underline{\beta}}_{WLS} = (X'V^{-1}X)^{-1}X'V^{-1}\underline{y} \qquad (1\text{-}9)$$

$\hat{\underline{\beta}}_{WLS}$ is an unbiased estimator for $\underline{\beta}$ (as is $\hat{\underline{\beta}}$). The covariance matrix for $\hat{\underline{\beta}}_{WLS}$ is (1-10)

$$\text{Cov}(\hat{\beta}_{WLS}) = (X'V^{-1}X)^{-1}\sigma^2 \qquad (1\text{-}11)$$

1-5. Previous Work on Factor Screening in Simulation

Although there is a substantial literature on factor screening, there has been little analysis or interpretation of this methodology in the computer simulation environment. Kleijnen [1975a,b], [1977] and Hunter and Naylor [1970] have suggested the use of fractional factorial designs and group screening (a procedure in which factors are arranged in sets) methods in simulation. However, they do not give any examples. Only Kleijnen [1975b] attempts to give any guidelines for the choice of a factor screening strategy. Nolan and Sovereign [1972] employ a group-screening strategy in a large-scale simulation model of airlift and sealift operations. However, they do not give any details of the screening methods used. Williams and Weeks [1974] have proposed using special types of p^n factorial designs for factor screening in simulation. Their methodology requires potentially many computer simulations runs, and there are no examples or evaluation of their methodology given. In general, there does not presently seem to be any systematic collection or evaluation of factor screening methods available, nor is there much specific analysis of their use in computer simulation. Some aspects of this will be dealt with in this report.

2. EXPERIMENTAL DESIGN METHODS IN FACTOR SCREENING

2-1. Full Factorial Designs

Full factorial experiments could be used for factor screening. The most efficient design to consider is the 2^k factorial; i.e., k factors each at two levels. It is relatively standard practice to denote the factors by upper case letters such as A, B, etc., rather than the x_1, x_2, etc. notation

used previously. The statistical model for a 2^k design would include k main effects, $\binom{k}{2}$ two-factor interactions, $\binom{k}{3}$ three-factor interactions, ..., one k-factor interaction. That is, for a 2^k design the complete model would contain $2^k - 1$ effects. Two systems of notation for treatment combinations are widely used. For example, in a 2^5 design abd denotes the treatment combination with factors A, B, and D at the high level and factors C and E at the low level. A system of + and - signs is also useful, occassionally, where + denotes the high level of a factor and - denotes the low level. Thus ++-+- and abd are equivalent notations. The treatment combinations may be written in standard order by introducing the factors one at a time; each new factor being successively combined with those above it. For example, the standard order for a 2^4 design is (1), a, b, ab, c, ac, bc, abc, d, ad, bd, abd, cd, acd, bcd, and abcd.

For even a moderate number of factors the total number of runs in a 2^k factorial design is large. For example, a 2^5 has 32 treatment combinations, a 2^6 has 64 treatment combinations, and so on. Since resources are usually limited, the number of replicates that the experimenter can employ may be restricted. Frequently, available resources will only allow a single replicate of the design to be run, unless the experimenter is willing to omit some of the original factors. Most factor screening experiments would fall into this category.

With only a single replicate of the 2^k it is impossible to compute an estimate of experimental error, that is, a mean square for error. Thus, it seems that hypotheses concerning main effects and interactions cannot be tested. However, the usual approach to the analysis of a single replicate of the 2^k is to assume that certain higher-order interactions are negligible. The statistical analysis of these designs is well-known (see John [1971] or Montgomery [1976]). Either Yates' tabular algorithm or the regression approach outlined in Section

1 may be used to estimate the effects. The variance of the estimate of any effect is $N^{-1}\sigma^2$, where N is the total number of observations, assuming that observations are independent. Note that the regression treatment of the data in Table 1 is the analysis of a 2^2 design. The smallest design for which this procedure is recommended is the 2^4.

The practice of combining higher-order interaction mean squares to estimate the error is subject to criticism on statistical grounds. If some of these interactions are significant, then the estimate of error will be inflated. As a result, other significant effects may not be detected and the significant interactions used as error will not be discovered. As a general rule, it is probably unwise to assume two-factor interactions to be zero without prior information. If most two-factor interactions are small, then it seems likely that all higher-order interactions will be significant also. (A word of caution here--one does not have to look very far for counterexamples to these rules).

In most factor screening studies, we will be willing to assume that certain high-order interactions (say three-factor and higher) are negligible. Considering the amount of information provided by a 2^k factorial, this is probably reasonable. For example, consider a 2^5. The 32 observations allow 31 effects to be estimated:

5 main effects
10 2 factor interactions
10 3 factor interactions
5 4 factor interactions
1 5 factor interactions

In many situations, out interest would be confined to detecting main effect and the 2-factor interactions. Thus we could either use the 16 higher-order

effects as an estimate of error, or as the basis of developing a more efficient design via fractional replication.

When a large number of effects are estimated, we may wish to find some formal basis for declaring which effects are significant. If there is either replication or insignificant factors pooled to estimate error, we could possibly use analysis of variance methods and conduct formal statistical tests. However, if variance reduction methods such as common random numbers have been used, the usual analysis of variance statistical tests may not be appropriate. For a discussion of this problem is simple designs, see Heikes, Montgomery, and Rardin [1976]. A useful approach is to plot the effects on normal probability paper. Negligible effects on such a display will fall approximately along a straight line, while real effects will lie far from the line. For examples of this methodology in a general experimental design setting, see Montgomery [1976]. We will illustrate the approach in subsequent examples.

The 2^k factorial series has a projection property useful in factor screening. For example, consider the 2^3 design in Figure 3. If factor A is negligible, we can collapse the 8 runs from the 2^3 in factors A, B, and C into two replicates of a 2^2 in factors B and C. In general, if we have a single replicate of a 2^k and $h(<k)$ factors can be dropped because they seem negligible, then the remaining data will always correspond to 2^h replicates of a full factorial in the remaining k-h factors. These replicated design points can be used to obtain an estimate of error.

Full 2^k factorial are advantageous in screening in that they potentially produce all of the information required to identify significant effect and interactions. However, there are more resource-efficient methods that can produce equivalent information.

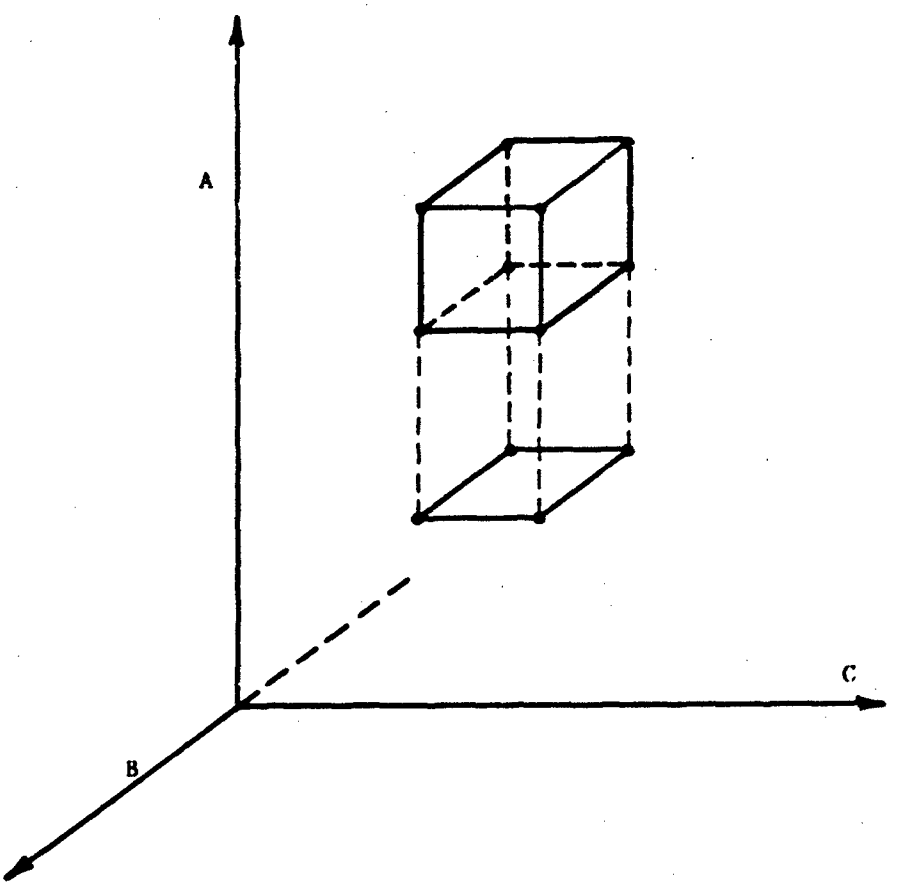

Figure 3. Projection of a 2^3 into a 2^2 Design in the Factors B and C

2-2. The 2^{k-p} Fractional Factorial Design

2-2.1 General Results

As the number of factors in a 2^k factorial design increases, the number of runs required for a complete replicate of the design rapidly outgrows the resources of most experimenters. A complete replicate of the 2^6 design requires 64 runs. In this design only 6 of the 63 degrees of freedom correspond to main effects, and only 15 degrees of freedom correspond to two-factor interactions. The remaining 42 degrees of freedom are associated with three-factor and higher interactions.

If the experimenter can reasonably assume that certain high-order interactions are negligible, then information on main effects and low-order interactions may be obtained by running only a fraction of the complete factorial experiment. These fractional factorial designs are widely used in industrial research, and have major applications in factor screening. For a general introduction to the construction and elementary properties of these designs refer to Montgomery [1976, ch. 10] or Box and Hunter [1961].

In a 2^{k-p} fractional factorial design, only a fraction of the 2^k treatment combinations are actually run. Specifically, a fraction of the 2^k design containing 2^{k-p} runs is called a $1/2^p$ fraction of the 2^k, or, more simply, a 2^{k-p} fractional factorial design. The designs discussed in this section are regular fractions, that is, estimates of the effects are orthogonal. The effects may be estimated by Yates' algorithm (John [1976], Daniel [1977], Montgomery [1976]) or by generating the contrast for any factor using the table of + and - signs for that design (which is equivalent to the regression approach outlined in Section 1). The variance of the estimate of any effect is $2^{p-k}\sigma^2$.

There are several methods of constructing these designs. One method of constructing a 2^{k-p} fractional factorial design is to select p independent

generators (no chosen generator is a generalized interaction of the others), constructing the 2^p blocks associated with those generators, and then selecting one block as the fractional design. The defining relation for the design consists of the p generators initially chosen and their $2^p - p - 1$ generalized interactions.

The alias structure may be found by multiplying each effect modulus 2 by the defining relation. Care should be exercised in choosing the generators so that effects of potential interest are not aliased with each other. Each effect has $2^p - 1$ aliases. In most factor screening studies we assume higher-order interactions (say third- or fourth-order and higher) to be negligible, and this greatly simplifies the alias structure.

A second method of design construction is to consider the 2^{k-p} design as a full factorial in h = k-p factors. Then the table of + and - signs for the full 2^k design is written down, and the additional p factors added by equating their factor levels with the products of certain factor levels in the full 2^k. As an example, consider the 2^{6-2} design. This is a 1/4 fraction of a 2^6, containing $2^{6-2} = 2^4 = 16$ rows. To construct this design form a 2^4 design in the factors A, B, C, and D, as shown in the left-hand panel of Table 3. Two columns must be added to incorporate the fifth and sixth factors, E and F. These factor levels are found in the center panel of Table 3, by equating E = ABC and F = ACD. Note that this is equivalent to choosing generators I = ABCE and I = ACOF and using the first procedure described above to construct the design. The treatment combinations are shown in the right-hand panel of Table 3.

Since the generators of this design are I = ABCE and I = ACDF and the generalized interaction of the generators ABCE and ACDF is BDEF, the complete defining relation for this design is I = ABCE = ACDF = BDEF. To find the aliases of any effect multiply that effect by each word in the defining relation.

Table 3

Construction of the 2^{6-2} Design With Generators I = ABCE and I = ACDF

A	B	C	D	E = ABC	F = ACD	Treatment Combination
-	-	-	-	-	-	(1)
+	-	-	-	+	+	aef
-	+	-	-	+	-	be
+	+	-	-	-	+	abf
-	-	+	-	+	+	cef
+	-	+	-	-	-	ac
-	+	+	-	-	+	bcf
+	+	+	-	+	-	abce
-	-	-	+	-	+	df
+	-	-	+	+	-	ade
-	+	-	+	+	+	bdef
+	+	-	+	-	-	abd
-	-	+	+	+	-	cde
+	-	+	+	-	+	acdf
-	+	+	+	-	-	bcd
+	+	+	+	+	+	abcdef

For example, the alias of A is

$$A = BCE = CDF = ABDEF$$

It is easy to verify that every main effect is aliased by three-factor and five-factor interactions, while two-factor interactions are aliased with each other and with higher-order interactions. Thus, when we estimate A, for example, we are really estimating A + BCD + CDF + ABDEF. The complete alias structure is shown in Table 4. If three-factor and higher interactions are negligible, this design gives clear estimates of main effects.

The 2^{k-p} fractional factorial design has the projection property noted previously for the full 2^k design. In general, say 2^{k-p} fractional factorial design can be projected into either a full factorial or a replicated fractional factorial in some subset of $r = k-p$ of the original factors. Those subsets of factors providing fractional factorials are subsets appearing as words in the complete defining relation. This is particularly useful in screening experiments, when we suspect at the outset of the experiment that most of the original factors will have small effects. The original 2^{k-p} fractional factorial can then be projected into a full factorial (say) in the most interesting factors.

For example, the 2^{6-2} fractional factorial will collapse to a single replicate of a 2^4 design in any subset of four factors that is not a word in the defining relation. It will also collapse to a replicated one-half fraction of a 2^4 in any subset of four factors that is a word in the defining relation. Thus, the design in Table 3 becomes two replicates of a 2^{4-1} in the factors ABCE, ACDF, and BDEF, since these are the words in the defining relation. There are 12 other combinations of the six factors, such as ABCD, ABCF, and so on, for which the design projects to a single replicate of the 2^4. This

Table 4

Alias Structure for the 2^{6-2} Design With I = ABCE = ACDF = BDEF

Effect	Alias		
A	BCE	CDF	ABDEF
B	ACE	DEF	ABCDF
C	ABE	ADF	BCDEF
D	ACF	BEF	ABCDE
E	ABC	BDF	ACDEF
F	ACD	BDE	ABCEF
AB	CE	BCDF	ADEF
AC	BE	DF	ABCDEF
AD	CF	BCDE	ABDF
AE	BC	CDEF	ABDE
AF	CD	BCEF	ABDE
BD	EF	ACDE	ABCF
BF	DE	ABCD	ACEF
ABF	CEF	BCD	ADE
CDE	ABD	AEF	CBF

design will also collapse to two replicates of a 2^3 in any subset of three of the six factors or four replicates of a 2^2 in any subset of two factors.

To present a fractional factorial for which the projection property can be visually demonstrated, co sider the 1/2 fraction of the 2^3 with generating relation I = ABC. This could also be denoted as a 2^{3-1} design. The design is shown in Table 5. The projection of this design into a full 2^2 factorial is accomplished by eliminating one of the original three factors. This is illustrated in Figure 4.

2-2.2 Resolution III Designs

It is useful to classify 2^{k-p} frac'ional factorial designs according to their resolution. The system is as follows:

(i) Resolution III Designs. These are designs in which no main effects is aliased with any other main effect, but main effects are aliased with two-factor interactions and two-factor interactions are aliased with each other. The 2^{3-1} design in Table 5 is of resolution III.

(ii) Resolution IV Designs. These are designs in which no main effect is aliased with any other main effect or two-factor interaction, but two-factor interactions are aliased with other. The 2^{4-1} design with I = ABCD is of resolution IV.

(iii) Resolution V Designs. These are designs in which no main effect or two-factor interaction is aliased with any other main effect or two-factor interaction, but two-factor interactions are aliased with three-factor interactions. A 2^{5-1} design with I = ABCDE is of resolution V.

In general, the resolution of a design is equal to the smallest number of letters in any word in the defining relation. Consequently some authors refer to these plans as three-letter, four-letter, and five-letter designs, respectively.

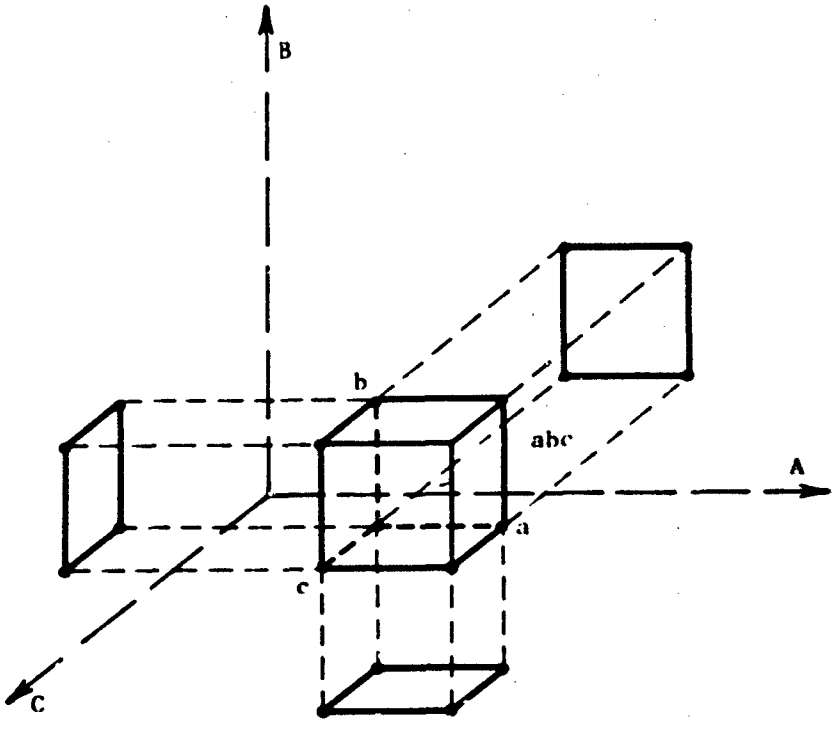

Figure 4. Projection of a 2^{3-1} Design Into a 2^2 Design

Table 5

The 2^{3-1} Design With I = ABC

A	B	C=AB	Treatment Combinations
-	-	+	c
+	-	-	a
-	+	-	b
+	+	+	abc

We can show that a design is of resolution (2t+1) if we can estimate effects of order t when effects of order higher than t are negligible. Roman numeral subscripts are used to identify the resolution of a design. Thus, a 2^{3-1}_{III} design is a 2^{3-1} design of resolution III. For the more highly fractionated designs, more extensive assumptions are required to draw conclusions from the data.

Resolution III and IV designs are particularly useful in factor screening studies. This section will discuss the 2^{k-p}_{III} design. We may construct resolution III designs for investigating up to k = N - 1 factors in N runs, where N is a multiple of 4. Designs in which N is a power of 2 can be constructed by the methods presented previously. Of particular importance are designs requiring 4 runs for up to 3 factors, 8 runs for up to 7 factors, 16 runs for up to 15 factors, and 32 runs for up to 31 factors. If k = N - 1 the fractional factorial design is said to be saturated.

A design for analyzing up to three factors in four runs is the 2^{3-1}_{III} design, presented in Table 5. Another very useful saturated fractional factorial is a design for studying seven factors in eight runs; that is, the 2^{7-4}_{III} design. This design is a one-sixteenth fraction of the 2^7. It may be constructed by first writing down the plus and minus levels for a full 2^3 in A, B, and C, and

then generating the levels of four additional factors using the interactions of the original three as follows: D = AB, E = AC, F = BC, and G = ABC. Thus, the generating relations for this design are I = ABD, I = ACE, I = BCF, and I = ABCF. The design is shown in Table 6.

Table 6

The 2^{7-4}_{III} Design With Generators I = ABD, I = ACE, I = BCF, and I = ABCF

A	B	C	D=AB	E=AC	F=BC	G=ABC	
-	-	-	+	+	+	-	def
+	-	-	-	-	+	+	afg
-	+	-	-	+	-	+	beg
+	+	-	+	-	-	-	abd
-	-	+	+	-	-	+	cdg
+	-	+	-	+	-	-	ace
-	+	+	-	-	+	-	bcf
+	+	+	+	+	+	+	abcdefg

The complete defining relation for this design is

$$I = ABD = ACE = BCF = ABCG = BCDE = ACDF = CDG = ABEF = BEF = AFG = DEF = ADEG = CEFG = BDFG = ABCDEFG$$

To find the alias of any effect multiply that effect by each word in the defining relation. For example, the alias of B is

$$B = AD = ABCE = CF = ACG = CDE = ABCDF = BCDG = AEF = EG = ABFG = BDEF = ABDEG = BCEFG = DFG = ACDEFG$$

This design is a one-sixteenth fraction, and since the signs chosen for the generators are positive, this is the principal fraction. It is also of resolution III, since the smallest number of letters in any word of the defining contrast is three. Any one of the 16 different 2^{7-4}_{III} designs could be constructed by using the generators with one of the 16 possible arrangements of signs in $I = \pm ABD$, $I = \pm ACE$, $I = \pm BCF$, $I = \pm ABCG$. All of these designs would be said to belong to the same family.

The eight runs in this design may be used to estimate the seven main effects. These estimates are obtained as linear combinations of the observations, where the signs in a particular linear combination are given in the associated column of Table 6. Thus, to estimate A, use the plus and minus signs in the A column. Each of these effects has 15 aliases; however, if we assume that three-factor and higher interactions are negligible, then considerable simplification in the alias structure results. Making this assumption, each of the linear combinations

$$\begin{aligned} \ell_A &= A + BD + CE + FG \\ \ell_B &= B + AD + CF + EG \\ \ell_C &= C + AE + BF + DG \\ \ell_D &= D + AB + CG + EF \\ \ell_E &= E + AC + BG + DF \\ \ell_F &= F + BC + AG + DE \\ \ell_G &= G + CD + BE + AF \end{aligned} \qquad (2\text{-}1)$$

where ℓ_i refers to the linear combinations of treatment combinations given by column i in Table 6.

The saturated 2^{7-4}_{III} design in Table 6 can be used to obtain resolution III designs for studying fewer than seven factors in eight runs. For example,

to generate a design for six factors in eight runs, simply drop any one column in Table 6, for example, column G. This produces the design shown in Table 7.

Table 7

A 2^{6-3}_{III} Design With Generators I = ABD, I = ACE, and I = BCF

A	B	C	D=AB	E=AC	F=BC	
-	-	-	+	+	+	def
+	-	-	-	-	+	af
-	+	-	-	+	-	be
+	+	-	+	-	-	abd
-	-	+	+	-	-	cd
+	-	+	-	+	-	ace
-	+	+	-	-	+	bcf
+	+	+	+	+	+	abcdef

It is easy to verify that this is a 2^{6-3}_{III} design or a one-eighth fraction of the 2^6. The defining relation for the 2^{6-3}_{III} design is equal to the defining relation for the original 2^{7-4}_{III} design with any words containing the letter G deleted. Thus, the defining relation for this design is

$$I = ABD = ACE = BCF = BCDE = ACDF = ABEF = DEF$$

In general, when d factors are dropped to produce a new design, the new defining relation is obtained as those words in the original defining relation that do not contain any dropped letters. When constructing designs by this method, care must be taken to obtain the best design. If we drop columns B, D, F, and

G from Table 6, we obtain a design for three factors in eight runs, yet the treatment combinations correspond to two replicates of a 2^2. The experimenter would probably prefer to run a full 2^3 design in A, C, and E.

It is also possible to obtain a resolution III design for studying up to 15 factors in 16 runs. This saturated 2^{15-11}_{III} design can be generated by first writing down the 16 treatment combinations associated with a 2^4 in A, B, C, and D, and then equating 11 new factors with the 2, 3, and 4-factor interactions of the original 4. A similar procedure can be used for the 2^{31-26}_{III} design, which allows up to 31 factors to be studied in 32 runs.

By combining fractional factorial designs in which certain signs are switched, we can systematically isolate effects of potential interest. The alias structure for any fraction with the signs for one or more factors reversed is obtained by making changes of sign on the appropriate factors in the alias structure of the original fraction.

Consider the 2^{7-4}_{III} design in Table 6. Suppose that along with this principal fraction a second fractional design with the signs reversed in the column for factor D is also run. That is, the column D in the second fraction is

$$- + + - - + + -$$

The effects that may be estimated from the first fraction are shown in (2-1) and from the second fraction we obtain

$$\ell_A^* = A - BD + CE + FG$$

$$\ell_B^* = B - AD + CF + EG$$

$$\ell_C^* = C + AE + BF - DG$$

$$\ell_D^* = -D + AB + CG + EF \qquad (2\text{-}2)$$

$$\ell_E^* = E + AC + BG - DF$$

$$\ell_F^* = F + BC + AG - DE$$

$$\ell_G^* = G - CD + BE + AF$$

assuming that three-factor and higher interactions are insignificant. Now from the two linear combinations of effects $\frac{1}{2}(\ell_i + \ell_i^*)$ and $\frac{1}{2}(\ell_i - \ell_i^*)$ we obtain

i	From $\frac{1}{2}(\ell_i + \ell_i^*)$	From $\frac{1}{2}(\ell_i - \ell_i^*)$
A	A + CE + FG	BD
B	B + CF + EG	AD
C	C + AE + BF	DG
D	AB + CG + EF	D
E	E + AC + BG	DF
F	F + BC + AG	DE
G	G + BE + AF	CD

Thus we have isolated the main effect of D and all of its two-factor interactions. In general, if we add to a fractional factorial design of resolution III or higher a further fraction with the signs of a <u>single factor</u> reversed, then the combined design will provide estimates of the main effect

of that factor and its two-factor interactions.

Now suppose we add to any fractional factorial design a second fraction in which the signs for <u>all factors</u> are reversed. This procedure breaks the alias links between main effects and two-factor interactions. That is, we may use the <u>combined</u> design to estimate all main effects clear of any two-factor interactions. For example, suppose we added to the 2^{7-4}_{III} design in Table 6 the second fraction shown in Table 8.

Table 8

A 2^{7-4}_{III} Design With All Signs Switched

A	B	C	D=AB	E=AC	F=BC	G=ABC	
+	+	+	-	-	-	+	abcg
-	+	+	+	+	-	-	bcde
+	-	+	+	-	+	-	acdf
-	-	+	-	+	+	+	cefg
+	+	-	-	+	+	-	abef
-	+	-	+	-	+	+	bdfg
+	-	-	+	+	-	+	adeg
-	-	-	-	-	-	-	(1)

The effects that may be estimated from this fraction are

$$\ell^*_A = -A + BD + CE + FG$$

$$\ell^*_B = -B + AD + CF + EG$$

$$\ell^*_C = -C + AE + BF + DG$$

$$\ell^*_D = -D + AB + CG + EF$$

$$\ell_E^* = -E + AC + BG + DF$$

$$\ell_F^* = -F + BC + AG + DE$$

$$\ell_G^* = -G + CD + BE + AF$$

Upon combining the two fractions and forming the linear combinations $\frac{1}{2}(\ell_i + \ell_i^*)$ and $\frac{1}{2}(\ell_i - \ell_i^*)$, we obtain

i	From $\frac{1}{2}(\ell_i + \ell_i^*)$	From $\frac{1}{2}(\ell_i - \ell_i^*)$
A	BD + CE + FG	A
B	AD + CF + EG	B
C	AE + BF + DG	C
D	AB + CG + EF	D
E	AC + BG + DF	E
F	BC + AG + DE	F
G	CD + BE + AF	G

Therefore clear estimates of all main effects and the two-factor interaction alias groups are obtained.

The designs due to Plackett and Burman [1946] are also two-level Resolution III fractional factorials. These designs can be used for studying k = N - 1 variables in N runs, where N is a multiple of 4. If N is a power of 2, these designs are identical to those presented earlier in this section. However, for N = 12, 20, 24, 28, and 36 the Plackett-Burman designs are frequently useful.

The upper panel of Table 9 presents rows of plus and minus signs used to construct the Plackett-Burman designs for N = 12, 20, 24, and 36, while the

lower panel of the table presents blocks of plus and minus signs for constructing the design for N = 28. The designs for N = 12, 20, 24, and 36 are obtained by writing the appropriate row in Table 9 as a column. A second column is then generated from this first one by moving the elements of the column down one position and placing the last element in the first position. A third column is produced from the second similarly, and the process continued until column k is generated. A row of minus signs is then added, completing the design. For N = 28, the three blocks X, Y, and Z are arranged as

X Y Z

Z X Y

Y Z X

and a row of minus signs added to these 27 rows. The design for N = 12 runs and k = 11 is shown in Table 10.

The alias structure of the Plackett-Burman designs is complex. In general, all two-factor interactions not involving factor Q (say) are aliased with the estimate of Q. For example, in the 11 factor plan shown in Table 10, each main effect is aliased with 45 two-factor interactions, and each two-factor interaction appears in 9 of the 11 estimates of main effects. This is somewhat less troublesome if fewer than 11 factors are considered. Furthermore, the two-factor interactions could possibly be untangled by adding a second fraction with all signs reversed, provided that only a few of them were large.

EXAMPLE 1. We shall now illustrate some of the above ideas with an example. The problem setting is inventory control, and we wish to determine the effect of various parameters on the average annual cost. We note that simulation

Table 9

Plus and Minus Signs for the Plackett-Burman Designs

k = 11	N = 12 + + - + + + - - - + -
k = 19	N = 20 + + - - + + - + - + - + - - - - + + -
k = 35	N = 36 - + - + + + - - - + + + + + - + + + - - + - - - - + - + - + + - - + -

k = 27, N = 28

+ - + + + + - - -	- + - - - + - - +	+ + - + - + + - +
+ + - + + + - - -	- - + + - - + - -	- + + + + - + + -
- + + + + + - - -	+ - - - + - - + -	+ - + - + + - + +
- - - + - + + + +	- - + - + - - - +	+ - + + + - + - +
- - - + + - + + +	+ - - - - + + - -	+ + - - + + + + -
- - - - + + + + +	- + - + - - - + -	- + + + - + - + +
+ + + - - - + - +	- - + - - + - + -	+ - + + - + + + -
+ + + - - - + + -	+ - - + - - - - +	+ + - + + - - + +
+ + + - - - - + +	- + - - + - + - -	- + + - + + + - -

Table 10

Plackett-Burman Design for N = 12, k = 11

A	B	C	D	E	F	G	H	I	J	K
+	-	+	-	-	-	+	+	+	-	+
+	+	-	+	-	-	-	+	+	+	-
-	+	+	-	+	-	-	-	+	+	+
+	-	+	+	-	+	-	-	-	+	+
+	+	-	+	+	-	+	-	-	-	+
+	+	+	-	+	+	-	+	-	-	-
-	+	+	+	-	+	+	-	+	-	-
-	-	+	+	+	-	+	+	-	+	-
-	-	-	+	+	+	-	+	+	-	+
+	-	-	-	+	+	+	-	+	+	-
-	+	-	-	-	+	+	+	-	+	+
-	-	-	-	-	-	-	-	-	-	-

methods are not required for this problem, as there are analytical models that can be used to describe the system. However, the problem has been kept simple deliberately to illustrate the experimental methods.

There are three items in the inventory. These items are military belts, such as used in jeans and other casual apparel. Item 1 is hardware, item 2 is dyed webbing, and item 3 is natural webbing. The following quantities are fixed:

	Item 1	Item 2	Item 3
Annual Demand (D)	500,00 doz.	300,000 doz.	200,000 doz.
Demand during a Lead time $X \sim N(\mu,\sigma^2)$	$\mu_1 = 20{,}000$ $\sigma_1 = 3{,}000$	$\mu_2 = 6{,}000$ $\sigma_2 = 900$	$\mu_3 = 4{,}000$ $\sigma_3 = 600$
Lead Time τ	2 weeks	1 week	1 week
Fixed Cost A	\$35 per order	\$15 per order	\$15 per order
Unit Var. cost C	\$6.25/doz.	\$3.10/doz.	\$2.80/doz.
Carrying cost h	\$.20	\$.28	\$.28
Cost per unit short π	*	\$.40	\$.40

The following variables represent parameters that we would like to investigate to learn their effect on the system:

Variable	Level	Item 1	Item 2	Item 3
Order quantity Q	1	10,000	4,000	3,000
	2	20,000	8,000	6,500
Reorder point r	1	17,000	5,000	3,500
	2	35,000	11,000	7,000
Cost per unit π^* short	1	.3		
	2	.5		

Note that there are seven factors, each at two levels. The 2^{7-4}_{III} design in Table 6 is run, using the high and low levels of these factors shown above. Let factors A, B, and C denote the order quantities for items 1, 2, and 3; D, E, and F denote the reorder points for items 1, 2, and 3; and G denote the shortage cost for item 1. From the design in Table 6, we obtain the following:

Treatment Combination	Response $ X1000	Effect + Aliases (2-1)	Estimate
(def)	4,626	A	-65
afg	4,693	B	50
beg	4,718	D	-180
abd	4,655	C	-66
cdg	4,662	E	-72
ace	4,653	F	-58
bcf	4,685	G	80
abcdefg	4,626		

Obviously, the effect of D (and its aliases) is large. Since this is the only large effect, we might stop and conclude that over the range of variation, that only item 1's reorder point seriously affects the system. However, to be more certain of these results, we run the alternate fraction given in Table 8. This gives the following:

Treatment Combination	Response $ X1000	Effect + Aliases (2-2)	Estimate
abcg	4,683		
bcde	4,632	-A	66
acdf	4,656	-B	114
cefg	4,704	-D	182
abef	4,647	-C	-32
bdfg	4,640	-E	72
adeg	4,640	-F	24
(1)	4,716	-G	-16

Combining the results from the two fractions, we obtain

i	From $\frac{1}{2}(\ell_i + \ell_i^*)$	From $\frac{1}{2}(\ell_i - \ell_i^*)$
A	BD + CE + FG = 1	A = -65
B	$\underline{AD}$ + CF + EG = 82	B = -32
C	$\underline{AE}$ + BF + DG = -49	C = -17
D	AB + CG + EF = 1	D = -181
E	AC + BG + DF = 0	E = -72
F	BC + AG + $\underline{DE}$ = -17	F = -41
G	CD + BE + AF = 32	G = 48

Clearly the main effect of D is large. Since the effect of D is over twice as large as the next largest effect, we are tempted to conclude that it is the only significant factor. This is confirmed by viewing the normal probability plot of the estimates of the effects, Figure 5. Point 1 on this plot is D. It is significantly off the straight line formed by the other effects. We conclude that only the effect of D is significant.

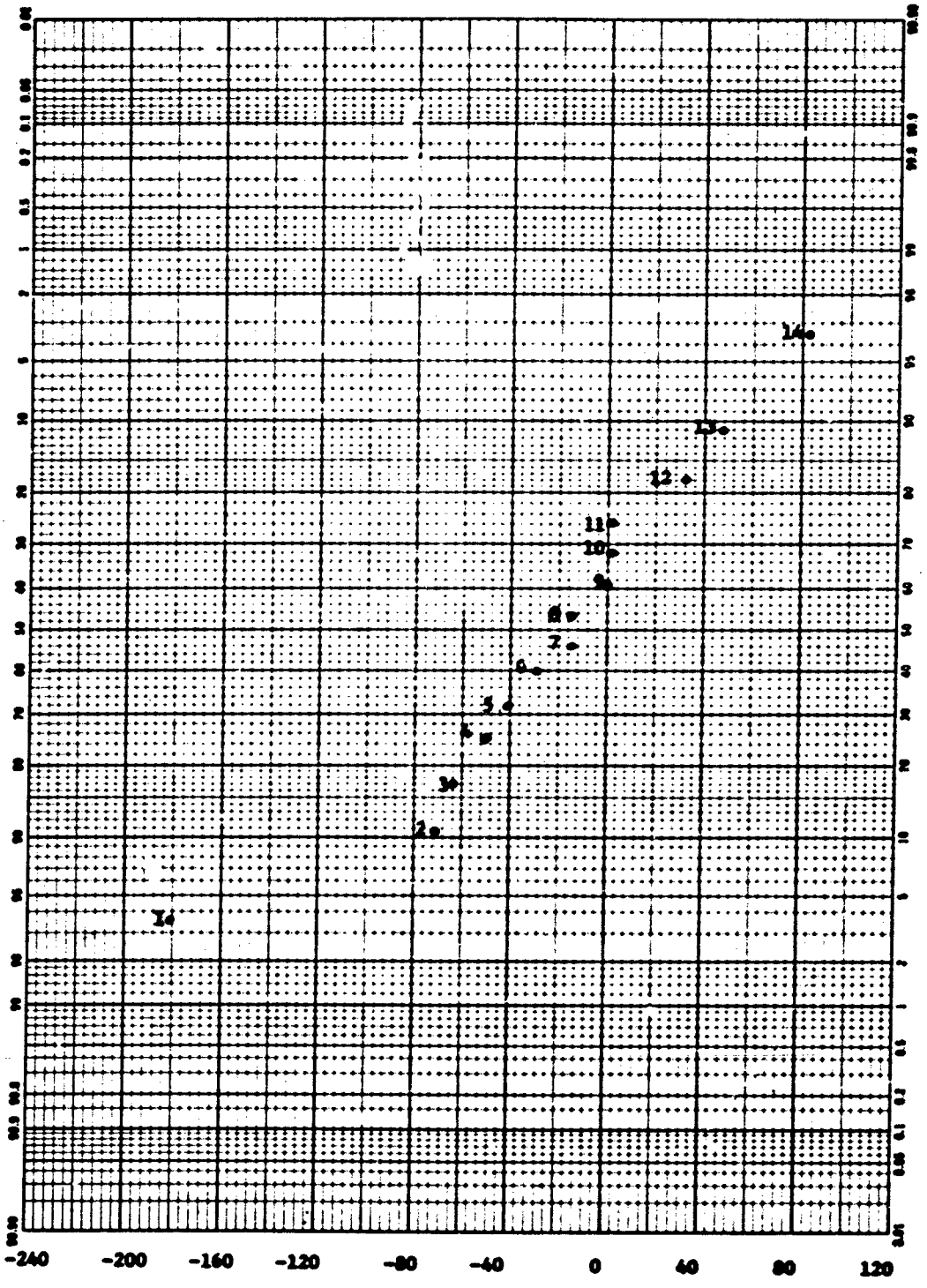

Figure 5. Normal Probability Plot, Example 1.

2-2.3 Resolution IV Designs

A 2^{k-p} fractional factorial design is of resolution IV if main effects are clear of two-factor interactions and some two-factor interactions are aliased with each other. Thus, if three-factor and higher interactions are suppressed, main effects may be estimated directly in a 2^{k-p}_{IV} design. The 2^{6-2} design in Table 3 is of resolution IV. Furthermore, the two combined fractions of the 2^{7-4}_{III} design in Example 1 is a 2^{7-3}_{IV} design.

Any 2^{k-p}_{IV} design must contain at least 2k runs. Resolution IV designs that contain exactly 2k runs are called <u>minimal</u> designs. Resolucion IV designs may be obtained from resolution III designs by the process of <u>fold over</u>. To fold over a 2^{k-p}_{III} design simply add to the original fraction a second fraction with all signs reversed. Then the plus signs in the identity column I in the first fraction are switched in the second fraction, and a $(k+1)^{st}$ factor associated with this column. The result is a 2^{k+1-p}_{IV} fractional factorial design. The process is demonstrated in Table 11 for the 2^{3-1}_{III} design. The resulting design is a 2^{4-1}_{IV} design with generating relation I = ABCD.

Table 11

A 2^{4-1}_{IV} Design Obtained by Fold Over

	D			
	I	A	B	C
Original 2^{3-1}_{III} I = ABC	+	-	-	+
	+	+	-	-
	+	-	+	-
	+	+	+	+
Second 2^{3-1}_{III} with signs switched	-	+	+	-
	-	-	+	+
	-	+	-	+
	-	-	-	-

As a second example of fold-over, consider the 2^{7-4}_{IV} design used in Example 1 (also see Table 6). By adding to the design the fraction in Table 8 and associating an 8^{th} H factor with the column I = + in Table 6 and I = - in Table 8, we would have a 2^{8-4}_{IV} plan. The generating relation for this design is

$$I = ABDH = ACEH = BCFH = ABCG = BCDE = ACDF = CDGH$$
$$= ABEF = BEGH = AFGH = DEFH = BDFG = ADEG = CEFG$$
$$= ABDEFGH.$$

The generator of the new design will consist of all generators from the old design that contain an even number of letters and all generators from the old design that contain an odd number of letters will have the new letter added.

Any resolution IV design will contain a 2^3 complete factorial design. That is, it will provide r replicates of a 2^7 design any 3 of the original factors, provided the design contain $r2^3$ points. Thus the 2^{8-4}_{IV} plan above provides two replicates of a 2^3 in any subset of 3 of the original 8 factors. This often has important applications in screening.

EXAMPLE 2. Consider the inventory problem in Example 1. We will fold over the original 2^{7-4}_{III} design in this example, giving a 2^{8-4}_{IV} plan, with the 8^{th} factor taken to be the mean of the lead time demand distribution for item 1. In the first fraction the mean is 20,000, while in the second it is 25,000. The following results are obtained.

Treatment Combination	Response ($ X 1000)	Estimate	Effect
def	4626		
afg	4693	96	FG + AH + BD + CE
beg	4718	258	EG + AD + BH + CF
abd	4655	288	AB + CG + DH + EF
cdg	4662	-170	DG + AE + CH + BF
ace	4653	-24	AC + BG + EH + DF
bdf	4685	-58	BC + AG + FH + DE
abcdefg	4626	-8	CD + AF + GH + BE
abcgh	4742	278	-H
bcdeh	4631	224	-A
acdfh	4655	158	-B
cefgh	4822	648	-D
abefh	4682	-38	-C
bdfgh	4639	120	-E
adeg	4639	58	-F
h	4786	-168	-G

Once again, only the effect of D appears large. This is confirmed by the normal probability plot given in Figure 6.

2-2.4 Remarks on Computations and Aliasing

To this point we have used the relatively simple computational methods associated with the 2^{k-p} designs, assuming that at least a regular fraction is available. Sometimes an experimenter will want to update the estimates of the effects following each additional run. This might often occur when augmenting a 2^{k-p} design with additional runs to estimate certain interactions. If we

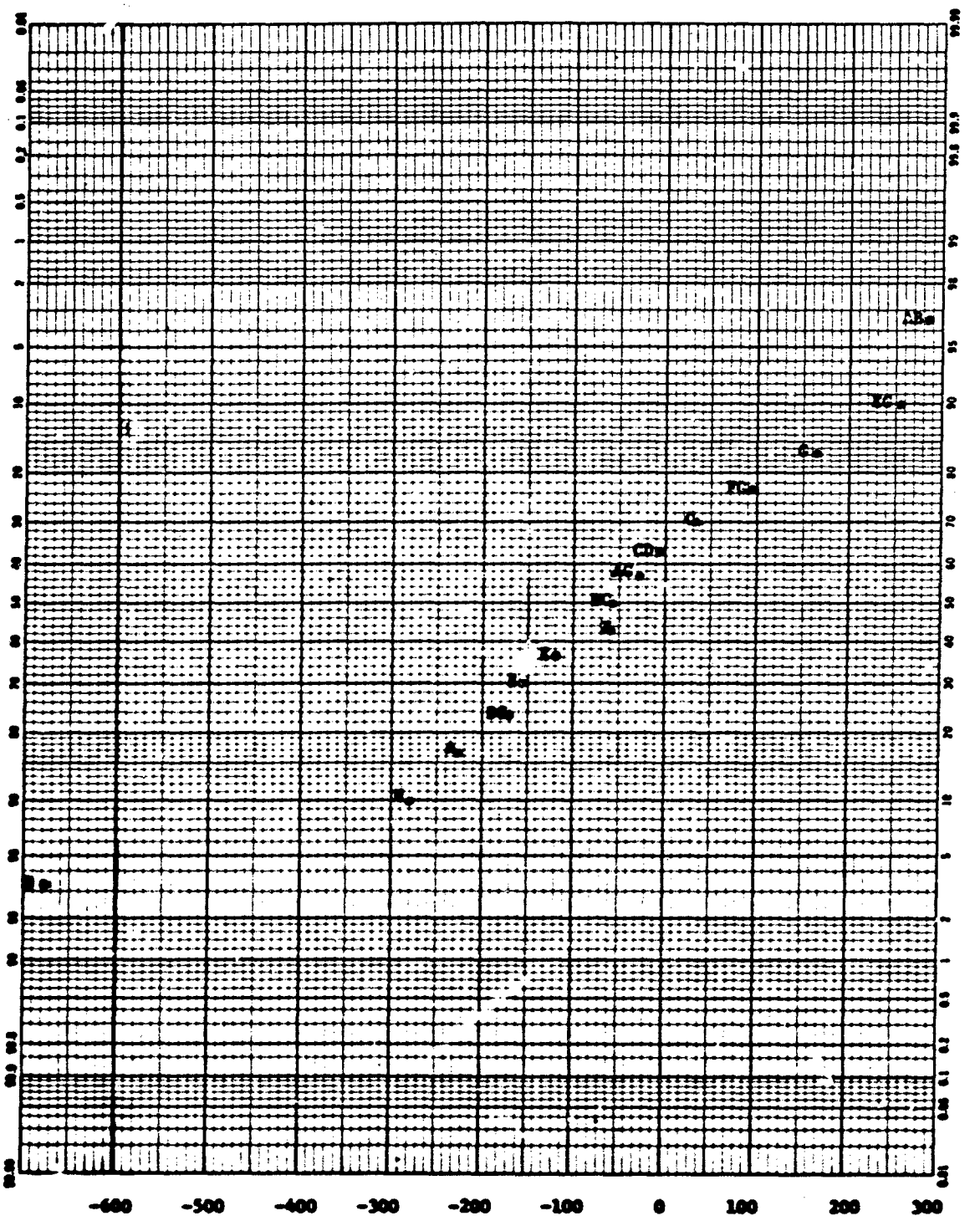

Figure 6. Normal Probability Plot, Example 2

consider the model

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon}$$

we can give an updating equation for $\hat{\underline{\beta}}$ in terms of <u>each</u> new run, assuming that the starting point was a block of runs giving orthogonal minimum variance estimates of $\underline{\beta}$ (such as the 2^{k-p} designs). This updating equation is

$$\hat{\underline{\beta}}_{NEW} = \hat{\underline{\beta}}_{OLD} + (MN + p)^{-1} \sum_{i=1}^{n} (y_i - \hat{y}_i)\underline{x}_i \qquad (2\text{-}3)$$

where p is the number of model parameters, N is the block size, m is the number of blocks completed, y_i is the new observation associated with the new vector of variable settings $\underline{x}_i (i=1,2,\ldots,n \leq N)$, and $y_i = \hat{\underline{\beta}}'_{OLD}\underline{x}_i$. Equation (2-3) was derived by Hunter [1964].

We may also give a general result concerning aliasing. If the true model is

$$\underline{y} = X_1\underline{\beta}_1 + X_2\underline{\beta}_2 + \underline{\varepsilon},$$

but the experimenter has estimated only the parameters $\underline{\beta}_1$ using the model

$$\underline{y} = X_1\underline{\beta}_1 + \underline{\varepsilon}$$

then it is well known that $\hat{\underline{\beta}}_1$ is biased, such that

$$E(\hat{\underline{\beta}}_1) = \underline{\beta}_1 + A\underline{\beta}_2$$

where the matrix $A = (X_1'X_1)^{-1}(X_1'X_2)$ is called the alias matrix. This general result can be used to work out the aliases for effects in the 2^{k-p} system. It is often useful in more complex design settings than the 2^{k-p}, particularly in irregular fractions, such as discussed in the next section.

2-3. Irregular Fractions of the 2^k Design

There are some multifactor screening situations in which higher saturation of the design than can be accomplished with regular fractions would be justified. This would be the case, for example, when computer runs are very time-consuming or expensive. In these situations, certain irregular fractional factorial designs may be useful. Often in these designs, the experimenter will only be able to estimate certain parameters in the model and will have few remaining degrees of freedom. Furthermore, the estimates of the effects will generally be nonorthogonal.

The simplest irregular fractions result from augmentation of a balanced 2^{k-p} fraction. One may view the process of combining fractions from the same family in the 2^{k-p}_{III} series as augmentation designs, where the augmented set is as large as the initial set. The methods presented here are based on smaller augmented sets, usually 1, 2, 4, or 8 runs, added with the objective of estimating two-factor interactions.

As an elementary example, consider the 2^{3-1}_{III} design. If only the A effect is large, then an estimate of the A effect clear of the BC interaction can be obtained with only one additional run. Thus if $I = -ABC$, and the runs made are (1), ab, ac, and bc. Now consider observation a. Since $E(a) = \mu + A - B - C - AB - AC + BC$, we have, if $B = C = AB = AC = 0$,

$$E(a) = \mu + A + BC$$

If we have an estimate $\hat{\mu}$ from the original fraction, then A + BC is estimated by $\ell^* = a - \hat{\mu}$. We can estimate A - BC directly from the first fraction as $\ell = -(1) + ab + ac - bc$. Then $\ell^* + \ell$ estimates A and $\ell^* - \ell$ estimates BC.

Similar augmentation schemes can be derived for most other designs in the 2^{k-p} series, either to separate a single two-factor interaction, a pair of two-factor interactions, or four such interactions. Daniel [1972] is the basic reference in this area. Addelman [1969] discusses the same problem, in more detail than Daniel [1962], but with less adaptation of results to special cases.

Three-quarter replicates of the 2^{k-p} series are often highly useful. These designs may be viewed as constructed by either omitting a quarter-fraction from the full 2^k or by adding a quarter-fraction to a one-half fraction. A good survey of these designs is in John [1971]. We will illustrate one of these designs with an example.

EXAMPLE 3. Suppose that in Example 1, only items 1 and 2 are of interest. We would like to obtain estimates of all 4 main effects (the order quantities and reorder points) and the 6 two-factor interactions. Obviously a 2^{4-1} will not do, since it contains only 8 runs and we must estimate 10 parameters. The full 2^4 design, requiring 16 rows, is considered too expensive. Only 12 rows can be taken.

We can estimate all 10 effects with 12 observations by using a 3/4 fraction of the 2^4. Consider the quarter replicates (2^{4-2}, $I = \pm AB = \pm ACD$):

(1) $I = +AB = +ACD = +BCD$; d, ab, c, abcd

(2) $I = +AB = -ACD = -BCD$; (1), abd, cd, abc

(3) $I = -AB = +ACD = -BCD$; bd, a, bc, acd

(4) $I = -AB = -ACD = +BCD$; b, ad, bcu, ac.

Omit the first fraction and run only the last three. Now overlap these three quarter replicates as follows to estimate the effects:

Fraction 1: (2) + (,) J = -BCD

$\underline{A}$ - ABCD = -110

$\underline{AB}$ - ACD = 0

$\underline{AD}$ - ABC = 0

ABD - $\underline{AC}$ = -110

Fraction 2: (2) + (4) I = -ACD

$\underline{B}$ - ABCD = -32

$\underline{AB}$ - BCD = 0

$\underline{BD}$ - ABC = 32

ABD - $\underline{BC}$ = 0

Fraction 3: (3) + (4) J = -AB

$\underline{C}$ - ABC = -318

$\underline{D}$ - ABD = -88

$\underline{CD}$ - ABCD = 0

The estimates of the 4-main effects and 6 two-factor interactions are shown above, assuming that higher-order interactions are negligible. Once again, note that only the reorder point for item 1 seems to produce a significant result.

Addelman and Kempthorne [1961] have developed a series of orthogonal main effect plans. These designs are useful in cases where only main effects are of interest. In many cases factors with either 2 or 3 levels can be considered. Much other work has been done on irregular fractions of the 2^k3^m series.

Margolin [1968] [1972] has done much of the work in this area. Webb [1965] [1971] has also developed very compact mixed fractional factorials from this series, involving 20 or fewer runs. These plans all have very heavy 2 factor interaction aliasing. Of related interest is Webb [1968].

2-4. Supersaturated Plans

These are two-level designs devised by Booth and Cox [1962]. In these designs, each of k factors appears at the high and low levels N/2 times, where $N \leq k$. We assume that N is even. Clearly not all estimates of the effects can be orthogonal, since $N \leq k$. Booth and Cox [1962] generated these designs to obtain "near-orthogonality" by using the design criterion

$$\min\left(\max_{i \neq j} |\underline{d}_i' \underline{d}_j|\right)$$

where $\underline{d}_i$ is a row vector denoting the levels of factor i. The vector $\underline{d}_i$ will consist of N/2 + 1's and N/2 - 1's. Booth and Cox [1962] tabulate designs for N = 12 and $k \leq 16, 20, 24$; N = 18 and $k \leq 24, 30, 36$; and N = 24 and $k \leq 30$. They describe an algorithm for generating other designs, although the procedure may be very inefficient.

EXAMPLE 4. To illustrate the use of a supersaturated design, consider the inventory problem in Example 1. We now add a fourth item to the inventory, with the following parameters:

$$D = 350{,}000,\ \sigma_4 = 2{,}000,\ A = \$25,\ C = \$4.30,\ h = \$0.45,\ \pi = \$0.50$$

The following 13 factors will be considered in a screening experiment:

Factor	High Level	Low Level
Q_1	10,000	20,000
Q_2	4,000	8,000
Q_3	3,000	6,500
Q_4	5,000	9,000
r_1	17,000	35,000
r_2	5,000	11,000
r_3	3,500	7,000
r_4	7,000	15,000
π_1	\$0.30	\$0.50
μ_1	20,000	25,000
μ_2	6,000	8,000
μ_3	2,500	5,500
μ_4	4,000	10,000

The 13 factor Booth and Cox design to investigate these factors, and the responses obtained, are shown below:

Run No.	Q_1 A	Q_2 B	Q_3 C	Q_4 D	r_1 E	r_2 F	r_3 G	r_4 H	u_1 I	u_1 J	u_2 K	u_3 L	u_4 M	Response
1	+	+	+	+	+	+	+	+	+	+	+	-	-	$6138
2	+	-	+	+	+	-	-	-	+	-	-	-	-	6166
3	-	+	+	+	-	-	-	+	-	-	+	+	+	6247
4	+	+	+	-	-	-	+	-	-	+	-	-	+	6310
5	+	+	-	-	-	+	-	-	+	-	+	+	+	6328
6	+	-	-	-	+	-	-	+	-	+	+	+	+	6275
7	-	-	-	+	-	-	+	-	+	+	+	+	-	6419
8	-	-	+	-	-	+	-	+	+	+	-	+	-	6358
9	-	+	-	-	+	-	+	+	+	-	-	+	+	6150
10	+	-	-	+	-	+	+	+	-	-	-	-	+	6158
11	-	-	+	-	+	+	+	-	-	-	+	-	-	6137
12	-	+	-	+	+	+	-	-	-	+	-	-	-	6135

The contrasts for each factor are obtained in the usual way. These contrasts are:

A = -71	H = -169
B = -205	I = 297
C = -109	J = 449
D = -295	K = 267
E = -819	L = 733
F = -197	M = 115

Clearly the largest factor effect is E (or r_1), followed closely by L (or u_3). There are also several other moderately large contrasts that may indicate

significant factors. This example illustrates one of the major disadvantages of a supersaturated design. Following the initial experiment, if several effects seem to be potentially active, there is no simple additional set of experiments that can be run to isolate the factors of interest. This is in contrast to the 2^{k-p} series, where additional fractions from the same family can always be used to gain further information on potentially active factors, or to untangle the interactions. Moreover, the aliasing that is present in the contrasts from a supersaturated design is very heavy and irregular, and this will frequently cause a confusing picture to the analyst. In this light, the supersaturated designs are likely to be little better than the "random balance" designs proposed by Satterthwaite [1959] and Budne [1959].

2-5. Group Screening Designs

2-5.1. General Approach

These designs are intended for use in situations where the following conditions apply:

1. The number of factors k is relatively large
2. All factors have the same prior probability of being active
3. There are no interactions between active factors
4. The direction of all effects is known
5. The errors associated with the observations are NID(0,σ^2).

A group screening design is conducted by forming the original k factors into g groups. Then each group is considered as a single factor and investigated through a design such as the 2^{g-p}. If a group-factor is negligible, then all factors within that group are considered insignificant. Group factors that exhibit significant effects are then divided into smaller groups for subsequent experimentation.

These designs were introduced by Watson [1961], who proposed that only two stages be used. Thus in the second stage, we experiment with the original factors. Patel [1962] and Li [1972] have generalized these results to multiple stages.

2-5.2 Two-stage Group Screening

The k factors will be divided into g groups. Watson [1961] originally suggested that all groups be of the same size, although this assumption is unnecessary. Because the direction of effects is known, we can label the high level of each factor as the level producing the largest response. The upper level of a group factor consists of running each factor in the group at the high level. If this arrangement is not followed, some factor effects may cancel.

Watson [1961] derives the optimum group size to be

$$f^* = [(1-\alpha_1)p]^{-1/2} \qquad (2\text{-}4)$$

where p is an estimate of the fraction of active factors and α_1 is the significance level used for the first-stage statistical analysis. This formula attempts to minimize the total number of runs required in both stages. It also implies that groups will be of equal size. If we have no prior estimate of p, or if the direction of some effects are not known, then (2-4) is invalid.

Generally, we would expect p to vary from factor to factor. That is, we would have considerable knowledge about some factors, and little knowledge about others. Note that as p increases, the optimum group size decreases. Therefore, it would seem reasonable to use groups of different sizes, depending on our knowledge of p for each factor. Factors that we strongly suspect are significant would be run in very small groups (perhaps of size 1). Furthermore, factors for which we do not know the direction of the effect could be tested in

groups of size 1 to prevent the cancellation effect.

As a hypothetical example of group screening, suppose we have 17 factors. Suppose that the direction of factor 1 is unknown, and that we are almost positive that factor 2 is active. The possible directions of the other is factors are known. Therefore, a logical arrangement of the groups would be:

Group Factors	Original Factors
A	1
B	2
C	3,4,5,6,7
D	8,9,10,11,12
E	13,14,15,16,17

These five factors could be investigated in the first stage using a 2^{5-2}_{III} design (8 runs). This would permit investigation of all main group effects, but these effects would be aliased with the two-factor interactions of the group effects. If we wanted to use 16 runs, the 2^{5-1}_{IV} design would allow estimation of all main effects and two-factor interactions of the group factors.

If the assumption of no active two-factor interactions between the original factors holds, then the factors may be formed into groups on an arbitrary basis. However, some choices of grouping arrangements will lead to more easily interpreted results, or to smaller sets of active factors to be investigated at the second stage. Sometimes we can use our knowledge of the problem to form the groups. For example, we might place all similar factors in the same groups. Thus if we are simulating an inventory system, all reorder quantities could form one group, all reorder levels a second group, etc. If some two-factor interactions may be active, then we must take more care in forming the groups. Generally, a significant two-factor interaction (say AB) biases the estimates of a third factor (say C) if and only if all three factors belong to separate

group factors. Therefore, if we suspect that some two-factor interactions are active, then all the factors involved in those interactions should be placed in the same group. For a proof of this result, see Kleijnen [1975a, b].

In the second stage of a group screening design, in addition to investigating the set of potentially active factors, we must also choose levels for the negligible factors identified in the first stage. Recall that the linear model can be written as

$$\underline{y} = X_1\underline{\beta}_1 + X_2\underline{\beta}_2 + \underline{\varepsilon},$$

where now $\underline{\beta}_1$ contains the set of potentially active factors and $\underline{\beta}_2$ contains the set of factors tentatively identified as negligible at the first stage. The matrix x_1 consists of the levels assigned to the active factors in the second stage and x_2 consists of the factors assigned to the negligible factors. Now, the expected value of the least squares estimate of $\underline{\beta}_1$, $\hat{\underline{\beta}}_1 = (X_1'X_1)^{-1}X_1'\underline{y}$, is

$$E(\hat{\underline{\beta}}_1) = \underline{\beta}_1 + (X_1'X_1)^{-1}X_1'X_2\underline{\beta}_2.$$

Clearly, if all the factors thought to be insignificant from stage 1 really are insignificant, then $\underline{\beta}_2 = \underline{0}$ and $\hat{\underline{\beta}}_1$ is an unbiased estimator of $\underline{\beta}_1$. However, if one or more of these effects is active, then $\underline{\beta}_2 \neq \underline{0}$ and $\hat{\underline{\beta}}_1$ is a biased estimator of $\underline{\beta}_1$.

The extent of the bias in $\hat{\underline{\beta}}_1$ is given by the alias matrix $A = (X_1'X_1)^{-1}X_1'X_2$. This may be controlled by the choice of factor levels for the variables in x_2. Assuming that two-level factors are employed, then if all levels in x_2 are identical (say +1, the high level) then the coefficients in $\underline{\beta}_2$ will bias only the intercept or overall mean term in $\underline{\beta}_1$. No other effects in $\underline{\beta}_1$ will be biased by factors in $\underline{\beta}_2$.

To prove this, suppose that X_1 is n x p, $\underline{\beta}_1$ is p x 1, X_2 is n x r, and $\underline{\beta}_2$ is r x 1. If the second-stage design is a 2^k or an orthogonal fraction of the 2^k, then $(X_1'X_1)^{-1} = (1/n)I_p$. Furthermore, if all of the negligible factors in X_2 are set at their high levels, then X_2 is an n x r matrix of 1's. Now X_1 is an n x p matrix, the first column of which consists of 1's (to account for the overall mean μ) and the remaining p-1 columns consist of the +1 and -1 levels from the orthogonal 2^k design. Therefore, $X_1'X_2$ is a p x r matrix, the first row of which consists of n's, all the remaining elemtns are all zero. Therefore,

$$(X_1'X_2)^{-1}X_1'X_2 = (1/n)I_p \begin{bmatrix} n & n & \cdots & n \\ \hline & & 0 & \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \hline & & 0 & \end{bmatrix},$$

and the alias structure is

$$E(\hat{\beta}_0) = \beta_0 + \sum_{i=p}^{r-1} \beta_i$$

$$E(\hat{\beta}_i) = \beta_i, \quad i=1,2,\ldots,p-1$$

Thus the r elements in $\underline{\beta}_2$ bias only the estimate of the intercept $\hat{\beta}_0$. Strictly speaking, all of the r negligible factors do not all have to be here at the high level. However, each factor must be held at the same level throughout the experiment.

EXAMPLE 5. Consider the inventory problem in Example 4. Suppose that there are 13 factors of interest, Q_1, r_1, μ_1, π_1, Q_2, r_2, μ_2, Q_3, μ_3, Q_4, r_4, and μ_4. We will arrange these factors in 4 groups, according to item, as follows:

Group Factor	Original Factor
A	Q_1, r_1, μ_1, π_1
B	Q_2, r_2, μ_2
C	Q_3, r_3, μ_3
D	Q_4, r_4, μ_4

A 2^{4-1} design is used to analyze these four group factors. The results are summarized below:

Treatment Combination	Response	Effect	Estimate
(1)	6207		
ad	6164	A + BCD	-180
bd	6183	B + ACD	-116
ab	6134	AB + CD	-10
cd	6210	C + ABD	6
ac	6168	AC + BD	4
bc	6181	BC + AD	-8
abcd	6135	D + ABC	2

Note that the two largest effects are A and B (and other aliases). Group factors C and D, and consequently the factors for item 3 and 4 are negligible. Therefore, following the initial 8 rows, we have reduced the set of potentially active factors from 13 to 7. The 7 remaining factors, Q, r_1, μ_1, π_1, Q_2, r_2, and μ_2 could be investigated using a 2^{7-4}_{III} or 2^{7-3}_{IV} plan, such as illustrated

earlier.

2-5.3 Group Screening With More Than Two Stages

Patel [1962] and Li [1962] have generalized Watson's results to more than two stages. Their procedures are very similar. Patel showed that the total number of runs is minimized if we choose the number of groups according to

$$g_1 \approx kp^{n/(n+1)}$$

$$g_2 = g_3 = \cdots = g_n = g_{n+1} \approx p^{-1/(n+1)},$$

where g_i is the number of groups into which each of the groups at stage i-1 is split. He also notes that an n-stage procedure is preferable to an n-1 stage procedure if

$$p < [1-(1/n)]^{n(n-1)}.$$

Group sizes decrease geometrically with parameter $p^{1(n+1)}$. Note that if we suspect that if more than one-fourth of the factors are active ($p > .25$), then the optimum number of stages is one. If between one-twelfth and one-fourth of the factors are active, then two stages should be used. Similarly, a three-stage procedure would be used if between one-thirtieth and one-twelfth of the factors are active. Clearly, these designs will be useful only in situations where p (the ratio of active to total factors) is thought to be very small.

2-6. Variance Reduction Considerations in Factor Screening

An important consideration in the design of a computer simulation experiment is the incorporation of variance reduction methods into the design. Two common variance reduction methods are the use of common pseudorandom numbers and

antithetic pseudorandom numbers for different points in the design. These methods have application to factor screening. Early work on this problem was by Fishman [1974]. Recently, a comprehensive treatment of the subject was published by Schruben and Margolin [1978].

We assume that when common random number streams are used at two design points, the two output statistics exhibit positive correlation, and when antithetic random number streams are used at any two points, negative correlation between outputs is induced. These assumptions are, of course, not always met in practice, but they are satisfied relatively often, as has been confirmed by numerous investigations (see Kleijnen [1975a], pp. 197-198).

Two possible estimation methods can be used, ordinary least squares (OLS), or weighted least squares (WLS). These estimators are

$$\hat{\underline{\beta}}_{OLS} = (X'X)^{-1}X'\underline{y}$$

and

$$\hat{\underline{\beta}}_{WLS} = (X'V^{-1}X)^{-1}X'V^{-1}\underline{y}$$

where V is the correlation matrix induced on the responses. The covariance matrices for these estimators are

$$\mathrm{Cov}(\hat{\underline{\beta}}_{OLS}) = (X'X)^{-1}X'V^{-1}X(X'X)^{-1}$$

and

$$\mathrm{Cov}(\hat{\beta}_{WLS}) = (X'V^{-1}X)^{-1}$$

A widely used criterion for comparing designs for estimating $\underline{\beta}$ is the determinant of the covariance matrix of the estimator. Designs that minimize this criterion are called D-optimal designs. The determinants of the covariance matrices associated with the OLS and WLS estimators are

$$D_{OLS} = |X'X|^{-2}|X'V^{-1}X|$$

and

$$D_{WLS} = |(X'V^{-1}X)^{-1}|$$

The WLS estimator has smallest generalized variance among the class of linear unbiased estimators. However, it is often impossible to calculate the WLS estimate because the matrix V is unknown.

There are some situations in which the OLS and WLS estimators are equivalent, and, hence, these two estimators would produce the same covariance matrix. Schruben and Margolin [1978] show that the two estimators are equivalent for the cases of the random number assignment schemes that minimize D_{WLS}. That is, an induced correlation structure that would minimize D_{WLS} is also one for which the estimators $\hat{\underline{\beta}}_{OLS}$ and $\hat{\underline{\beta}}_{WLS}$ (and have D_{OLS} and D_{WLS}) are identical. Therefore, the OLS estimator can be used.

Schruben and Margolin [1978] propose the following rule. If an experimental design admits orthogonal blocking into two blocks, then if for all points in block 1 we use the same common set of pseudorandom numbers, and for all points in block 2 we use the antithetic set of random numbers, then the OLS estimator of $\underline{\beta}$ will have minimum generalized variance. Specifically, this assignment rule will produce an estimator of β_0 that is superior to that obtained by common random numbers, and equivalent in terms of dispersion to common random numbers

for estimating the remaining parameters in β. In general, the best results are obtained if the block sizes are the same. Furthermore, the positive and negative correlations induced do not have to be equal.

There are some special results that can be stated for the 2^{k-p} series of designs. If the induced positive and negative correlations are equal in magnitude, then the assignment rule above produces a minimum generalized variance for the class of 2^{k-p} designs assuming that the linear model contains a mean (β_0) plus a subset of $r < 2^{k-p}$ effects. This assignment rule also minimizes the trace of the covariance matrix of $\hat{\underline{\beta}}$ (that is, the sum of the variances of $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_r$ is minimized).

Occasionally, factor screening experiments will make use of saturated designs. For a saturated design, <u>any</u> induced correlative structure between the observations results in an improvement with respect to the generalized variance criterion over that obtained from independently seeking each design point. Furthermore, the OLS and WLS estimators are equivalent in this case also.

These results have direct application to factor screening. Any 2^k or 2^{k-p} design that is not saturated can be run in two orthogonal blocks by identifying the blocks with the + and - levels of one of the k columns in the design. Thus, only k-1 factors could be investigated.

We now give some illustrations. First consider the 2^{6-2} design shown in Table 3. We can run this design in two blocks, say

block 1	
(1)	bdef
aef	abd
bcf	cde
abce	acdf

block 2	
be	df
abf	ade
cef	bcd
ac	abcdef

These blocks were formed by confounding the ABF effect (and its aliases; see Table 4) with blocks. The treatment combinations in block 1 would be run with one set of common random numbers and those in block 2 would be run with the antithetic set of random numbers.

As a second example, consider the 2^{7-4}_{III} design run in Example 1. Since 7 factors are considered in only 8 runs, this is a saturated fractional factorial. If only this fraction is to be run, any induction of correlation is superior to independent observations, so running all 8 observations with common random number streams would be an appropriate strategy. Now, if any fraction from the same family is added to the original fraction, the new fraction should be run using the antithetic random number stream. Clearly, this is an optimal strategy, since the two fractions together can be viewed as a fold-over design with the random number stream effect taking the levels of the eighth factor (which is + in the fraction 1 and - in fraction 2).

As a third example, consider the 2^{6-3}_{III} design in Table 7. This design investigates 6 factors in 8 runs, and since it is not a saturated fraction, we could obtain a minimum generalized variance by decomposing the design into two orthogonal blocks of 4 runs each. Now any nonsaturated Resolution III plan can be run in two blocks by identifying the + and - levels of a single additional variable with the blocks. Thus, in our example, add a seventh column to Table 7 by setting the signs in that column equal to 8 = ABC. Thus, the signs are -, +, +, -, +, -, -, and +. Consequently, run treatment combinations def, abd, ace, and bcf in block 1 (-) with a common set of random numbers, and treatment combinations a^c, be, cd, and abcdef in block 2 with the antithetic set of random numbers.

Now suppose upon examining the estimates of the effects from this fraction, it is decided to add a second fraction from the same family to separate main effects and two-factor interactions. The appropriate second fraction is

A	B	C	D=-AB	E=-AC	F=-BC	
+	+	+	-	-	-	abc
-	+	+	+	+	-	bcde
+	-	+	+	-	+	acdf
-	-	+	-	+	+	cef
+	+	-	-	+	+	abef
-	+	-	+	-	+	bdf
+	-	-	+	+	-	ade
-	-	-	-	-	-	(1)

In this new fraction, block 1 would consist of bcde, acdf, abef, and (1). These rows would be made with the same set of random numbers used in block 1 from the first fraction. Block 2 in the new fraction would consist of abc, cef, bdf, and ade. These runs would be made with the antithetic stream of random numbers used in block 2 in the original fraction. It is easy to verify that the final design is a 2^{6-2}_{IV} plan, with generation I=BCDE=ACDF=ABEF. The estimators from the combined design have minimum generalized variance.

2-7. Evaluation and Choice of Screening Designs

In this section, we will evaluate the characteristics of the various types of screening designs. Hopefully, this will provide guidance on the selection of designs in practice.

The 2^{k-p} fractional factorial design has many advantages in factor screening. If we can afford N runs, where N is a power of 2, Resolution III plans can be derived that incorporate up to N-1 factors. These plans require the experimenter to assume that two-factor and higher interactions are negligible. However, the assumptions regarding interactions can, to some extent, be checked by combining the original 2^{k-p}_{III} design with a second fraction from the same family.

If the experimenter can afford up to N=2k runs, the 2^{k-p} Resolution III and IV plans are highly recommended. The Plackett-Burman plans, also of Resolution III, are not generally recommended for factor screening unless the analyst knows in advance that all but a few two-factor interactions are negligible. The heavy aliasing of main effects and two-factor interactions is an undesirable property of these designs.

The supersaturated plans of Booth and Cox, like the Plackett-Burman designs, assume that only main effects are active. If this assumption is false, then the alias structure generated by a supersaturated design would be extremely difficult to untangle. The group screening methods of Watson and Patel are recommended instead. This approach would seem to have the economic efficiency required in simulation, without the overly-restrictive assumptions regarding interactions. For the vast majority of screening problems, either two or three stages will be sufficient. Once groups of factors are formed, it is recommended that 2^{k-p} fractional factorials be used to investigate the group factors.

3. SCREENING WITH UNDESIGNED AND PARTIALLY-DESIGNED DATA

3-1. Factor Screening with Regression Models

Very few factor screening studies will begin in an interactionless state. In most cases, we find that the analyst has some computational experience with the simulation model. It would be economically efficient to incorporate as much as possible of this historical information into the screening study.

In Section 2, we illustrated how the general linear model

$$\underline{y} = x\underline{\beta} + \underline{\varepsilon}$$

could be used in factor screening. If an experiment can be designed for studying the effect of the factors, very efficient parameter estimation techniques can be

used and data interpretation is relatively simple. One reason that the designed-experiment case is so simple is that most screening designs are orthogonal; that is, the regression coefficients $\hat{\underline{\beta}}$ have unconditional interpretations. If we apply the same approach to undesigned data that may have been collected for other purposes (such as validation or verification), this ease of interpretation is lost. However, it is still possible to learn something about the relative importance of the factors.

When dealing with undesigned or historical data, our approach is to fit an appropriate regression model to the data, and then make inferences on the model parameters to determine the effects of the factors. This is often hazardous, since it is well-known that the regression coefficients $\underline{\beta}$ measure only the partial effect of a variable. That is, β_j measures the effect of x_j conditional on the other x_i $(i \neq j)$ in the regression model. Furthermore, depending on the degree of nonorthogonality in the data, the least squares estimates of $\underline{\beta}$ may be very far from the true regression coefficients.

With undesigned data, the factor screening problem consists of two stages, (1) variable selection, and (2) interpretation of regression coefficients. We will discuss these problems in the next two sections.

There may also be a third type of screening study, part-way between the extremes of designed experiments and undesigned experiments. This is the situation in which some new data points may be collected for use with the original undesigned data, but the amount of new data to be added is not enough to constitute a fully-designed screening study. We will discuss methods for augmenting undesigned data for factor screening studies of this type.

3-2. Variable Selection Procedures

There is a vast literature on variable selection in regression. A very comprehensive review of this subject is in Hocking [1976]. Variable selection is

both an art and a science, and should be performed with care and caution. It should be regarded as exploration of the structure of the data.

We may classify variable selection methods into two general types, stepwise-type methods and search methods. Stepwise regression and its major variations (forward selection and backward elimination) are well-known. These procedures should not be used mechanically to find the "best" regression equation. Moreover, the order in which variables enter and leave the model should not be interpreted as measuring the relative importance of the factors. The existance of multicollinearity (correlation between factors), which is often a function of the disposition of the data in x-space, impacts the variable selection problem significantly.

Search-type variable selection methods include the all-possible regression algorithms, the Hocking-Lamotte SELECT procedure (see Hocking [1976] for a description), and the directed t-search method (see Daniel and Wood [1971]). These procedures often produce results superior to stepwise type methods, particularly for data that is badly nonorthogonal. The all possible regressions method has much to recommend it, particularly when the number of factors is small, say 20 or less. There are several good computationally efficient algorithms for all possible regressions, including the Furnival and Wilson [1974] algorithm, which is now available on BMD-P.

For factor screening purposes, stepwise type methods can be used at the outset of the problem, to reduce the number of factors to about 20. Generally, backward elimination seems to work well at this stage, although any of the stepwise-type procecures can produce good results if carefully used. Then one of the search methods such as all possible regressions, should be employed using the subset of the original factors identified at the first stage. The end result may be several final equations. Each good candidate equation should be examined for adequacy and validity using the standard techniques of residual analysis

(see Draper and Smith [1966], Ch. 3). Since the primary objective of building the regression model is to obtain good estimates of the parameters, the model selection criterion should be chosen accordingly. Selecting the model that gives a minimum mean square error will generally lead to good estimates of the individual regression coefficients. Selection of variables based on R^2 (a popular practice) often causes important variables to be left out of the equation.

3-1. Interpretation of Regression Coefficient

As noted previously, interpretation of regression coefficients is hazardous, since a_j measures the effect of x_j given that other factors x_i ($i \neq j$) are also in the model. Furthermore, the magnitudes of the individual coefficients are affected by the units of the factors and the response y. For this reason it is usually best to work with standardized coefficients (often identified as "beta coefficients" on regression computer program outputs). In general the standardized coefficients are found by solving

$$C\underline{a}^* = \underline{g} \tag{3-1}$$

where C is the correlation matrix of the k factors and $\underline{g}$ is a vector of simple correlations between x_j and the response y. The relationship between the standardized and original regression coefficients is

$$a_j = a_j^* \left(\frac{S_{yy}}{S_{jj}}\right)^{1/2}, \quad j=1,2,\ldots,k \tag{3-2}$$

where S_{yy} is the corrected sum of squares of y and S_{jj} is the corrected sum of squares of the x_j.

While the magnitude and sign of the standardized regression coefficients are often used as measures of importance of the factors, we must remember that

the partial nature of these coefficients still hampers interpretation. Only if the factors x_j, j=1,2,...,k are orthogonal (or nearly so) is the total effect of x_j reflected by β_j^*. Therefore, we should examine the final set of factors x_j, j=1,2,...,k, and measure the extent of departure from orthogonality before interpreting the individual standardized regression coefficients. One useful measure of orthogonality is $|C|$. If $|C| = 1$, the factors are orthogonal, while if $|C| = 0$, there is at least one linear dependency in the factors. Therefore, if $|C|$ is large, say close to 1, we feel relatively confident in interpreting the individual regression coefficients. On the other hand, if $|C|$ is small, say $|C| \leq 0.1$, then we suspect that severe multicollinearity is present, and, consequently, the regression coefficients are very unstable. In such a case, interpretation of the individual coefficients would be very risky.

For intermediate values, say $0.1 < |C| < 0.9$, other measures of multicollinearity should be examined. These include the variance inflation factors (the main diagonal elements of C^{-1}), and the eigenvalues of C. If the largest variance inflation factor is greater than 10, or if the ratio of the largest to smallest eigenvalue (called the conditioning number) exceeds 10, then corrective action should be taken before interpreting the individual coefficients. Such corrective action would consist of re-estimating the parameters by a method specifically designed to combat multicollinearity.

A widely-used parameter estimation method designed to combat multicollinearity is ridge regression. The ridge regression estimates are defined as

$$\underline{\beta}^*(k) = (C + kI)^{-1}\underline{g} \qquad (3\text{-}3)$$

where $k \geq 0$. The usual method for choosing k is to solve (3-3) for various k, plot $\underline{\beta}^*(k)$ versus k, and select k as the value at which reasonable stablization in the coefficients $\underline{\beta}^*(k)$ results. This plot is called the ridge trace. For

further details, see Hoerl and Kennard [1970].

If (3-2) is applied to the full set of factors, then the ridge trace is used to eliminate negligible factors. The rules for elimination of factors are:

1. Eliminate factors whose standardized coefficients are stable but small.

2. Eliminate factors whose coefficients are unstable and tend to zero as k increases.

3. Eliminate one or more factors with unstable coefficients.

The remaining set of variables should be examined for near-orthogonality. This may be done graphically by plotting $D = \hat{\beta}^*(k)'\hat{\beta}^*(k)$ against k. Note that D is the squared distance of $\hat{\beta}^*(k)$ from the origin. It can be shown that for an orthogonal system, the distance of the ridge coefficients from the origin should be $\hat{\beta}^*(0)'\hat{\beta}^*(0)/(1+k)^2$. If the factors are nearly orthogonal, the graph of these two functions should be nearly identical.

EXAMPLE 6. Consider the four-item inventory problem described earlier. Table 12 contains 30 observations on average annual cost and the corresponding values of the independent variables Q_1, Q_2, Q_3, Q_4, r_1, r_2, r_3, r_4, λ_1, and μ_1. These 30 runs do not correspond to any standard factor screening design. We will illustrate how regression methods can be used to identify the most influential factors.

These 10 variables were analyzed using the BMD-P stepwise multiple regression program P2R. The F-level for entering and removing variables was arbitrarily set at 4.0 (the logic for this choice stems from the fact that $t^2 = F$, and $t = 2.0$ corresponds roughly to 95 percent significance). The results of this analysis are summarized below:

Table 12

Data for the Inventory Problem, Example 6

Observation	Cost	Q_1	Q_2	Q_3	Q_4	r_1	r_2	r_3	r_4	Π_1	π_1
1	4449	10490	4230	5310	8970	22380	8810	3940	14200	.70	32010
2	6225	22370	3750	5350	7520	24130	8240	6800	21830	.27	39120
3	6181	9960	7790	5690	10490	37570	4100	6560	18740	.15	23490
4	6162	9600	8960	6230	6990	28920	11690	7510	7060	.39	27190
5	6194	8960	8540	4480	7390	32600	5150	4270	5680	.60	25830
6	6188	10360	6350	6120	10590	29300	10280	5860	8430	.18	35640
7	6160	17960	5180	3600	4540	29700	7070	3460	17620	.71	11400
8	6165	9220	5220	3580	11400	25900	9920	3080	9210	.57	22900
9	6174	14600	4860	5870	4880	17960	6200	3510	13600	.38	17030
10	6438	15010	5410	3990	4270	17430	4570	4230	4700	.56	25770
11	6198	9300	3260	6160	9140	18580	6990	4320	15400	.18	24010
12	6214	23940	8150	4440	9760	26410	9770	4320	6120	.36	34070
13	6176	18580	6900	6540	7250	20590	7250	4220	14800	.67	21990
14	6169	11000	4650	4080	3640	33060	10060	3050	8200	.47	16390
15	6492	8120	4690	4580	9620	27000	4700	3410	6140	.60	34480
16	6132	12550	3990	5000	5230	32360	10330	6040	9900	.55	26170
17	6218	21690	7280	4660	9800	21380	5100	6180	20500	.19	33330
18	6138	13780	3680	3580	9140	36860	9880	6470	21120	.83	17500
19	6385	20200	6960	5820	5150	15010	11150	3930	9260	.77	28100
20	6283	16440	4090	5920	4210	25590	6400	7560	7130	.90	29200
21	6287	19510	6110	5190	3580	22530	9870	4530	4900	.15	34590
22	6189	18060	5400	4940	3790	16660	4560	6260	21060	.76	12340
23	6714	9390	8540	6260	4930	19230	4870	3000	19930	.65	34900
24	6138	15110	6160	3780	8880	23490	11200	7900	17420	.40	15660
25	6195	21780	4270	3000	11160	26110	5550	5060	15020	.41	28270
26	6141	18230	7000	6950	6990	27080	10990	3920	8130	.84	10980
27	6162	17450	5390	6170	11500	28910	6120	6980	7660	.55	22220
28	6276	15900	3320	3940	8220	30670	7080	4400	10410	.50	39150
29	6138	13670	8560	3840	3790	39980	9190	6130	12240	.45	25000
30	6135	19120	6100	3980	7110	34150	8030	7620	14240	.13	24900

Variable	Standardized Coefficient	Partial F Statistic
Q_1	-0.372	7.438
r_1	-0.442	13.139
Π_1	0.378	8.940
μ_1	0.612	23.866

This equation has $R^2 = 0.6630$ and $MS_E = 6936.97$. A plot of residuals from this model versus the predicted values $\hat{y}_i$ is shown in Figure 7. This display indicates a tendency to underpredict cost near the extreme value of the response variable. This could occur either because important variables have been left out of the model, or because the relationship between cost and the independent variables is not linear. In this problem, considering that we know that average annual inventory cost is not a linear function of Q and r, it would seem that the latter possibility should be explored.

Since none of the variables associated with items 2, 3, or 4 are apparently significant, they are ignored, and the data analyzed with the candidate variables Q_1, r_1, Q_1r_1, r_1^2, Π_1, and μ_1. This second analysis is performed with the BMD-P all possible regressions algorithm (Furnival and Wilson [1974]) P9R. The criterion for model selection is minimum MS_E. The results are shown below:

Variable	Standardized Coefficient	t-statistic
Q_1	-1.350	-2.20
r_1	-2.995	-2.57
Q_1r_1	1.175	1.72
r_1^2	1.767	1.92
Π_1	0.401	3.34
μ_1	0.667	5.46

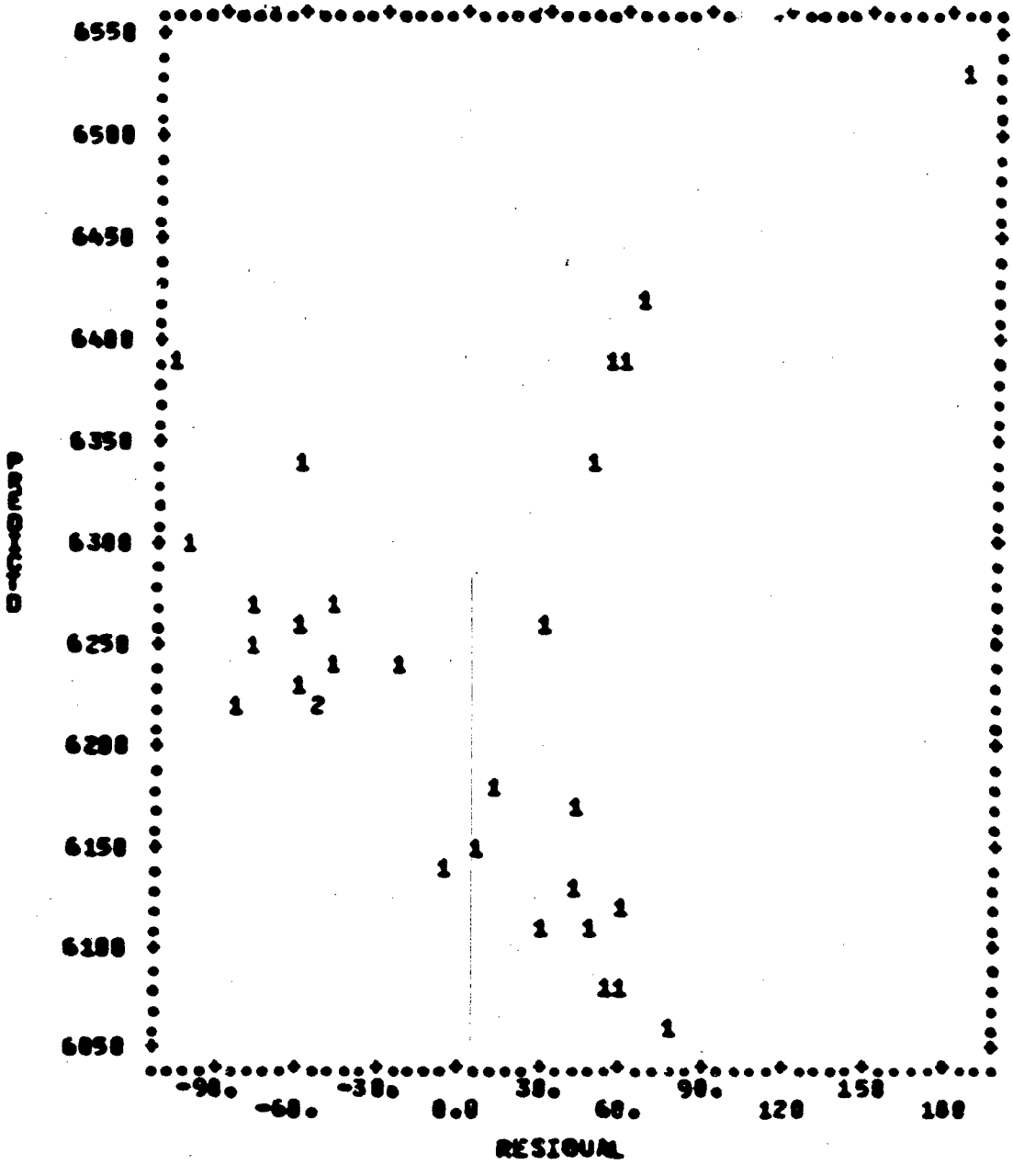

Figure 7. Plot of Residuals versus $\hat{y}_i$.

This [illegible] yields $R^2 = 0.7226$ and $MS_E = 6206.86$. Clearly Q_1 has a strong [illegible] effect, and r_1 exhibits both linear and second-order effects. The variables Q_1 and r_1 are much more influential than Π_1 and μ_1 in explaining the variation in average annual cost. There is also evidence of an interaction between Q_1 and r_1.

A plot of the residuals from this model versus the corresponding fitted values is shown in Figure 8, and a normal probability plot of residuals is shown in Figure 9. These displays do not indicate any gross violation of assumptions.

3-4. Augmenting Undesigned Data

In screening situations where some additional runs can be added to existing data, a natural question is the development of criteria for locating these new observations. If multicollinearity is a significant problem in the original data, then it seems logical to locate the new points so as to alleviate this problem, insofar as that is possible. On the other hand, if multicollinearity is not present, then other criteria could be developed.

A symptom of multicollinearity is a small value of $|C|$. Therefore, if m new runs are to be made, they should be at points in the factor space chosen to maximize $|C|$. If there are k factors, and if we think of the region of interest for these factors as a k-dimensional hypercube, then $|C|$ is maximized by adding m new runs at the corners of the experimental region. For details of this procedure, see Gaylor and Merrill [1968] and Dykstra [1966]. Their procedure allows the coordinates of all m new points to be determined simultaneously. If sequential augmentation is desired, then adding each new run at that point in the factor space where the variance of the predicted response is maximized will also maximize $|C|$.

Maximizing $|C|$ is a variance-oriented criterion. It is a reasonable criterion if the form of the model fit to the data is correct. However, in most factor screening studies, we have made the assumption that some effects are

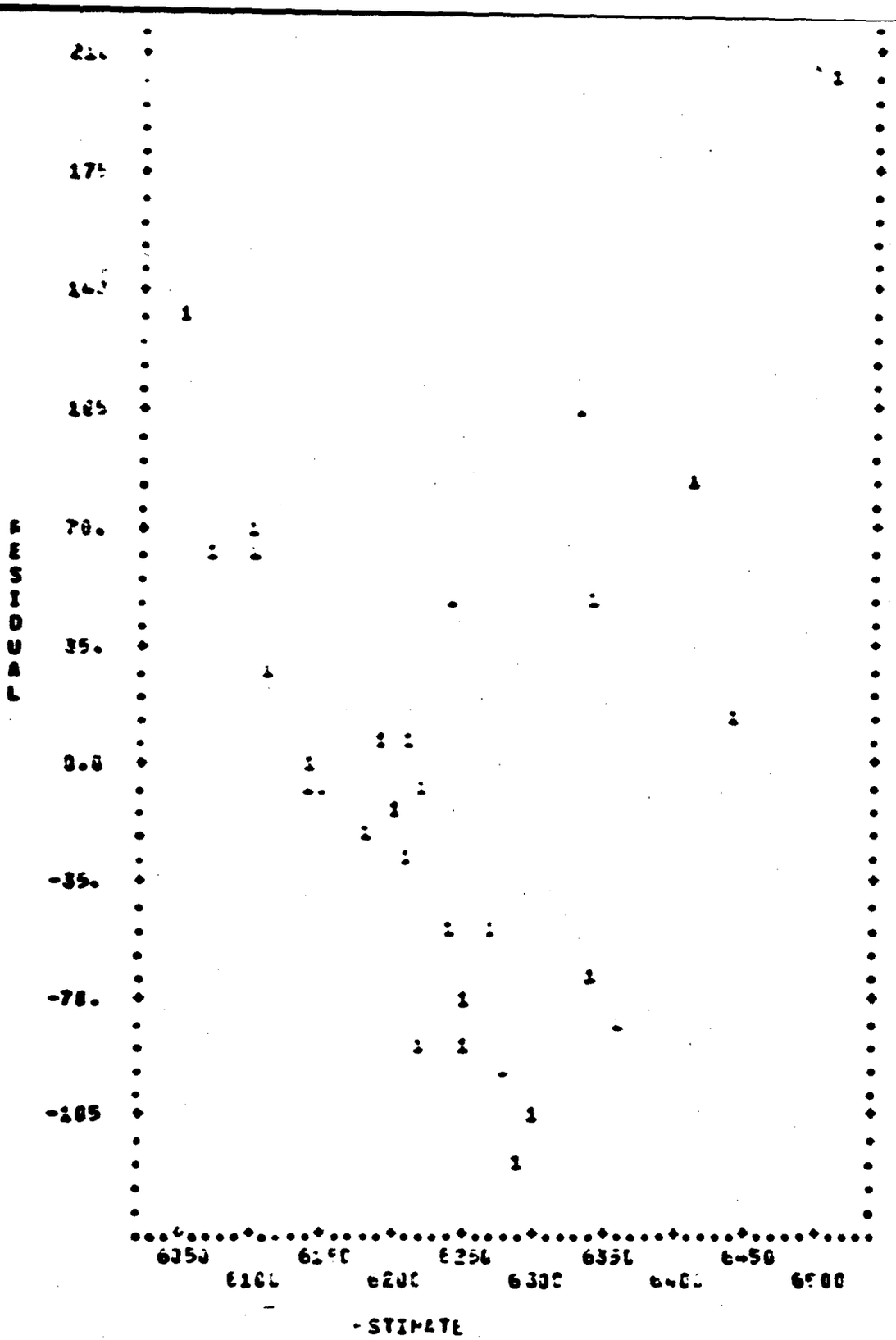

Figure 8. Plot of Residuals Versus Fitted Values, Example 6.

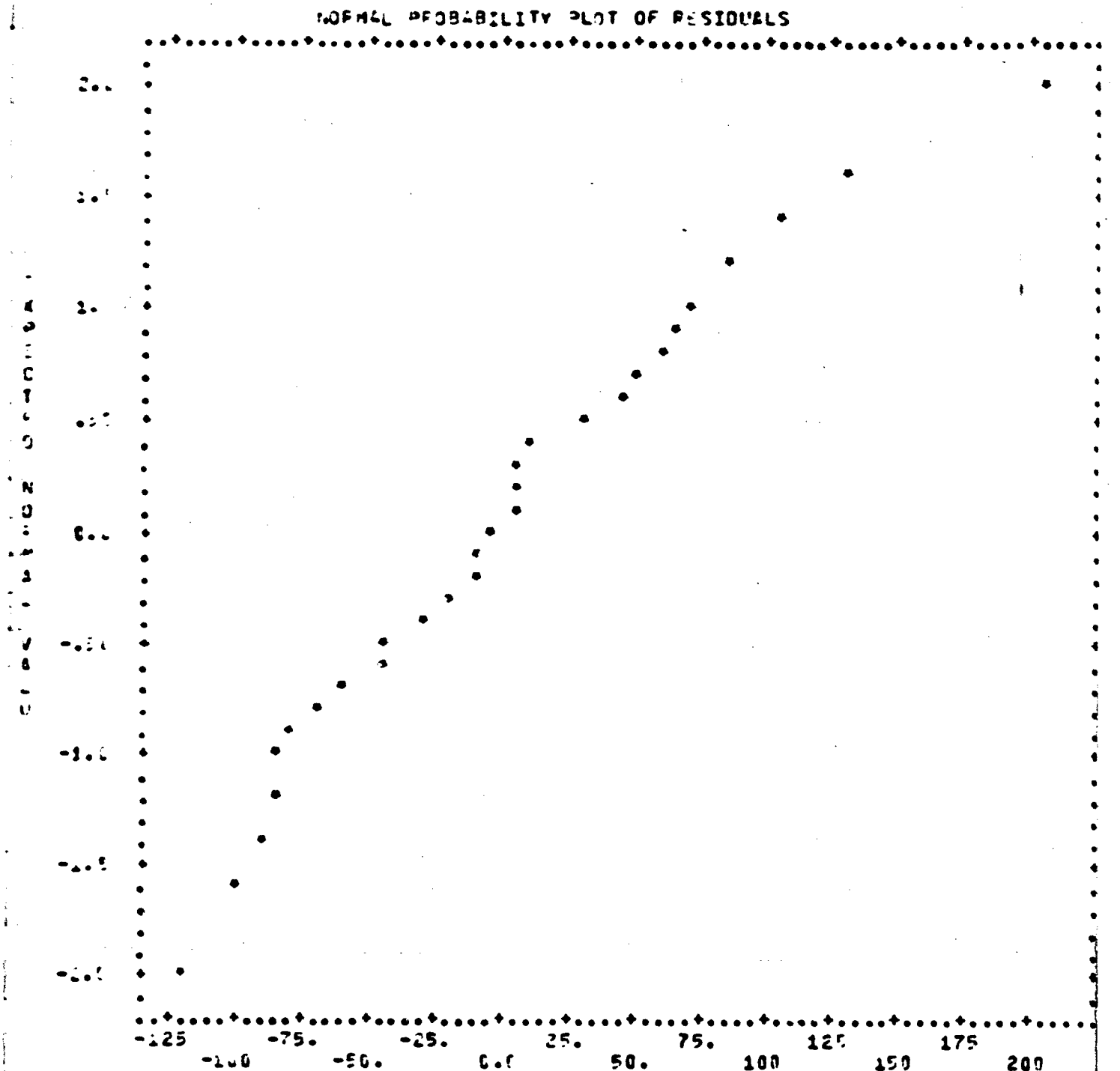

Figure 9. Normal Probability Plot of Residuals, Example 6.

negligible. Since there is always the possibility that these assumptions are incorrect, the analyst could elect to augment his original data with points chosen so that the bias in regression estimates from excluded factors is minimized. We will now describe a data augmentation scheme for this situation.

Suppose that we have fit the model

$$\underline{Y} = X_1 \underline{\beta}_1 + \underline{\epsilon}$$

but the response is really determined by the relationship

$$E(\underline{Y}) = X_1 \underline{\beta}_1 + X_2 \underline{\beta}_2.$$

Assume that the independent variables are defined such that the center of the region of interest R is at (0, 0, ... 0). The region of interest is a k-dimensional unit sphere and should include all points in the undesigned data. Care must be taken in selecting R since bias is not invariant under the transformation and different results would be obtained for different regions of interest.

The mean square error is a measure of both bias and variance. The mean square error, integrated over the region of interest, is

$$J = \int_R E[\hat{Y}(\underline{x}) - \eta(\underline{x})]^2 \, d\underline{x}$$

$$= \text{trace}[\mu_{11} M_{11}^{-1}] + \underline{\alpha}_2' [\mu_{22} - \mu_{12}' \mu_{11}^{-1} \mu_{12}]$$

$$+ (M_{11}^{-1} M_{12} - \mu_{11}^{-1} \mu_{12}) \mu_{11} (M_{11}^{-1} M_{12} - \mu_{11}^{-1} \mu_{12})] \underline{\alpha}_2 \qquad (3\text{-}4)$$

where $M_{ij} = N^{-1} X_i' X_j$, i, j=1,2, are matrices of design moments,

$$\mu_{ij} = \int_R \underline{X}_i \underline{X}_j \, dx \Big/ \int_R d\underline{x} \qquad (3\text{-}5)$$

are region moment matrices, $\underline{\alpha}_2 = \sqrt{N}\,\underline{\beta}_2/\sigma$, N is the number of observations, and σ^2 is the experimental error variance.

The average mean square error is composed of two terms, the average variance

$$V = \text{trace}[\mu_{11} M_{11}^{-1}], \qquad (3\text{-}6)$$

and the average squared bias

$$B = \underline{\alpha}_2' [(\mu_{22} - \mu_{12}\mu_{11}^{-1}\mu_{12}) + (M_{11}^{-1}M_{12} - \mu_{11}^{-1}\mu_{12})\mu_{11}(M_{11}^{-1}M_{12} - \mu_{11}^{-1}\mu_{12})]\underline{\alpha}_2. \qquad (3\text{-}7)$$

Average squared bias is minimized when design moments are equal to region moments, or

$$M_{11} = \mu_{11} \text{ and } M_{12} = \mu_{12} \qquad (3\text{-}8)$$

Average squared bias then is a function only of the region moments which are not controllable by the experimenter and its minimum value is

$$B_{min} = \underline{\alpha}_2' [\mu_{22} - \mu_{12}\mu_{11}^{-1}\mu_{12}]\underline{\alpha}_2 \, . \qquad (3\text{-}9)$$

An undesigned experiment will not meet the conditions in Equation (3-8), but it is possible to augment the experiment in such a way that the conditions will be met or nearly met. We then are operating on the controllable part of average squared bias, say

$$B_c = \underline{\alpha}_2' [(M_{11}^{-1}M_{12} - \mu_{11}^{-1}\mu_{12})\mu_{11}(M_{12}^{-1}M_{12} - \mu_{11}^{-1}\mu_{12})]\underline{\alpha}_2. \qquad (3\text{-}10)$$

Consider first the case of fitting a model containing all second order effects when some third order effects are present. We desire design moments to equal region moments through order 5. Expressing the equalities in equation (3-8) results in a set of simultaneous non-linear equations that can be solved for the additional experimental trials necessary. For example, the pure second design moments should equal the corresponding region moments, or

$$\sum_{u=1}^{N} x_{iu}^2 = N/(k+2), \quad i=1,\ldots,k, \qquad (3\text{-}11)$$

where k is the number of factors. Similar equations are written for the other moments. For the N observations already taken, the left hand side of Equation (3-11) is constant. We can now select m additional runs so that

$$\sum_{u=1}^{W} x_{iu}^2 = W/(k+2),$$

where W = N + m. The levels of the variables for the additional runs are x_{iu}, u=N+1,...,W.

The selection of the m additional runs is accomplished by minimizing the function

$$F(\underline{x}) = \sum_{i=1}^{k}\left(\sum_{u=1}^{W} x_{iu}\right)^2 + \sum_{i=1}^{k}\sum_{j>i}^{k}\left(\sum_{u=1}^{W} x_{iu}x_{ju}\right)^2$$

$$+ \sum_{i=1}^{k}\left(\sum_{u=1}^{W} x_{iu}^2 - W/(k+2)\right)^2 + \sum_{i=1}^{k}\sum_{j\geq i}^{k}\sum_{\ell\geq j}^{k}\left(\sum_{u=1}^{W} x_{iu}x_{ju}x_{\ell u}\right)^2$$

$$+ \sum_{i=1}^{k}\sum_{j\neq i}^{k}\left(\sum_{u=1}^{W} x_{iu}^3x_{ju}\right)^2 + \sum_{i=1}^{k}\sum_{j=i}^{k}\sum_{\substack{\ell\neq i\\ \ell>j}}^{k}\left(\sum_{u=1}^{W} x_{iu}^2x_{ju}x_{\ell u}\right)^2$$

$$+ \sum_{i=1}^{k} \sum_{j\neq i}^{k} \sum_{\ell>j}^{k} \sum_{p>\ell}^{k} \left(\sum_{u=1}^{W} x_{iu}x_{ju}x_{\ell u}x_{pu} \right)^2$$

$$+ \sum_{i=1}^{k} \sum_{j>i}^{k} \left(\sum_{u=1}^{W} x_{iu}^2 x_{ju}^2 - W/(k+2)(k+4) \right)^2$$

$$+ \sum_{i=1}^{k} \left(\sum_{u=1}^{W} x_{iu}^4 - W/(k+2)(k+4) \right)^2$$

$$+ \sum_{i=1}^{k} \sum_{j\geq 1}^{k} \sum_{\ell\geq j}^{k} \sum_{p\geq \ell}^{k} \sum_{q\geq p}^{k} \left(\sum_{u=1}^{W} x_{iu}x_{ju}x_{\ell u}x_{pu}x_{qu} \right)^2 \qquad (3\text{-}12)$$

For the case where we fit a main effects model and the true system contains second-order effects, moments through order 3 must be equal and this is accomplished by minimization of $F(\underline{x})$ in Equation (3-12) considering only the first four terms.

The bias will be minimized if the additional design points can be selected so that $F(\underline{x})$ is zero. However, in many cases when adding only a limited number of design points the minimum value of $F(\underline{x})$ is greater than zero, that is, not all moments can simultaneously be adjusted to the required values. In those cases where the minimum possible value of $F(\underline{x}) > 0$ the augmented design will minimize bias only when the contribution to the controllable part of bias resulting from any design moment not equalling the corresponding region moment is the same for all moments, that is, all components of $\underline{\alpha}_2$ are equal.

The value of $\underline{\alpha}_2$ increases as we add observations causing the minimum value of bias, B_{min}, to increase while the value of the controllable component of bias, B_c is decreasing. B_c can be decreased to zero, in which case further additional observations can only increase average squared bias due to the increase in B_{min}. Also the amount of increase in B_{min} may become greater than the decrease possible in B_c. This indicates that a measure is necessary that

will indicate when average squared bias is at a minimum.

We select as such a measure the <u>percentage</u> reduction in bias, say

$$PR = \frac{B_0 - B_a}{B_0} \times 100\%$$

where the subscript 0 indicates the original value of the undesigned experiment and the subscript a indicates the value after augmentation. By letting the terms inside the square brackets in equation 4 be denoted by Q, we can express PR as

$$PR = \frac{\underline{\alpha}'_{20} Q_0 \underline{\alpha}'_{20} - \underline{\alpha}'_{2a} Q_a \underline{\alpha}_{2a}}{\underline{\alpha}'_{20} Q_0 \underline{\alpha}_{20}}$$

$$= \frac{\sqrt{N_0}\,(\underline{\beta}_2/\sigma) Q_0 \sqrt{N_0}\,(\underline{\beta}_2/\sigma) - \sqrt{N_a}\,(\underline{\beta}_2/\sigma) Q_a \sqrt{N_a}\,(\underline{\beta}_2/\sigma)}{\sqrt{N_0}\,(\underline{\beta}_2/\sigma) Q_0 \sqrt{N_0}\,(\underline{\beta}_2/\sigma)}$$

or

$$PR = \frac{N_0 Q_0 - N_a Q_a}{N_0 Q_0} \qquad (3\text{-}13)$$

It can be seen that PR does not depend on the unknown value of $\underline{\beta}_2/\sigma$. The procedure to minimize bias is to determine the maximum number of new runs allowed, then sequentially select one run at a time by minimizing $F(\underline{x})$ m times. Any unconstrained search technique could be used to minimize $F(\underline{x})$.

REFERENCES

1. Addelman, S. [1969], "Sequences of Two-Level Fractional Factorial Plans", Technometrics, Vol. 11, No. 3, pp. 477-509.

2. Addelman, S. and O. Kempthorne [1961], Orthogonal Main-Effect Plans, ASTIA Arlington Hall Station.

3. Booth, K. H. Y. and Cox, D. R. [1962], "Some Systematic Supersaturated Designs", Technometrics, Vol. 4, No. 4, pp. 489-495.

4. Box, G. E. P. and Hunter, J. S. [1961], "The 2^{k-p} Fractional Factorial Designs", Parts I and II, Technometrics, Vol. 3, pp. 311-352, pp. 449-458.

5. Budne, T. A. [1959], "The Application of Random Balance Designs, Technometrics, Vol. 1, No. 2, pp. 139-155.

6. Burdick, D. S. and T. H. Naylor [1966], "Design of Computer Simulation Experiments for Industrial Systems", Communications of the ACM, Vol. 9, No. 8.

7. Cochran, W. G. and G. M. Cox [1957], Experimental Designs, Second Edition, John Wiley & Sons, New York.

8. Daniel, C. [1962], "Sequences of Fractional Replicates in the 2^{p-q} Series", Journal of the American Statistical Association, Vol. 58, No. 246, pp. 403-429.

9. Daniel, C. [1973], "One at a Time Plans", Journal of the American Statistical Association, Vol. 68, No. 285, pp. 353-360.

10. Daniel, C. [1976], Applications of Statistics to Industrial Experimentation, John Wiley & Sons, New York.

11. Daniel, C. and F. S. Wood [1971], Fitting Equations to Data, John Wiley & Sons, Inc., New York.

12. Davies, O. L. [1956], The Design and Analysis of Industrial Experiments, Hafner, New York.

13. Dykstra, O., Jr. [1966], "The Orthogonalization of Undesigned Experiments", Technometrics, Vol. 8, No. 2, pp. 279-290.

14. Fishman, G. S. [1973], Concepts and Methods in Discrete Event Digital Simulation, John Wiley & Sons, New York.

15. Fishman, G. S. [1974], "Correlated Simulation Experiments", Simulation, Vol. 23, pp. 177-180.

16. Furnival, G. M. and R. W. Wilson, Jr. [1974], "Regressions by Leaps and Bounds", Technometrics, Vol. 16, pp. 49[illegible].

17. Gaylor, D. W. and J. A. Merrill [1968], "Augmenting Existing Data in Multiple Regression", Technometrics, Vol. 10, No. 1, pp. 73-81.

18. Heikes, R. G., D. C. Montgomery, and R. L. Rardin [1976], "Using Common Random Numbers in Simulation Experiments - An Approach to Statistical Analysis", Simulation, Vol. 26, No. 3, pp. 81-85.

19. Hicks, C. R. [1973], Fundamental Concepts in the Design of Experiments, Second Edition, Holt, Rinehart, and Winston, New York.

20. Hocking, R. R. [1976], "The Analysis and Selection of Variables in Linear Regression", Biometrics, Vol. 32, pp. 1-49.

21. Hoevl, A. E. and R. W. Kennard [1970], "Ridge Regression: Application to Nonorthogonal Problems", Technometrics, Vol. 12, pp. [illegible]-93.

22. Hunter, J. S. [1964], "Sequential Fact[illegible]ion", Technometrics, Vol 6, No. 1, pp. 41-55.

23. Hunter, J. S. and T. H. Naylor [1970], "Experimental Designs for Computer Simulation Experiments", Management Science, Vol. 16, No. 7.

24. Ingall, E. J. [1972], "On Experimental Designs for Computer Simulation Experiments", Management Science, Vol. 18, pp. 384-388.

25. Jacoby, J. E. and S. Harrison [1962], "Multivariable Experimentati[illegible] and Simulation Models", Naval Research Logistics Quarterly, Vol. 9, pp. [illegible]21-136.

26. John, P. M. W. [1971], Statistical Design and Analysis of Experiments, McMillan, New York.

27. Kleijnen, J. P. C. [1975a], Statistical Techniques in Simulation, Parts I and II, Marcel Dekker, New York.

28. Kleijnen, J. P. C. [1975b], "Screening Designs for Poly-Factor Experimentations", Technometrics, Vol. 17, No. 4, pp. 487-493.

29. Kleijnen, J. P. C. [1977], "Design and Analysis of Simulations: Practical Statistical Considerations", Simulation, Vol. 27, No. 9, pp. 81-90.

30. Li, C. H. [1962], "A Sequential Method for Screening Experimental Variables", Journal of the American Statistical Association, Vol. 57, pp. 455-457.

31. Margolin, B. H. [1968], "Orthogonal Main Effect 2^n3^m Designs and Two-Factor Interaction Aliasing", Technometrics, Vol. 10, No. 3, pp. 559-573.

32. Margolin, B. H. [1972], "Non-Orthogonal Main-Effect Designs", Journal of the Royal Statistical Society, B, Vol. 34, No. 3, pp. 431-440.

33. Montgomery, D. C. [1976], Design and Analysis of Experiments, John Wiley & Sons, New York.

34. Montgomery, D. C. and D. M. Evans [1975], "Second Order Response Surface Designs in Computer Simulation, Simulation, Vol. 25, No. 6.

35. Naylor, T. H. (ed.) [1968], The Design of Computer Simulation Experiments, Duke University Press, Durham.

36. Nolan, R. L. and M. G. Sovereign [1972], "A Recursive Optimization and Simulation Approach to Analysis with an Application to Transportation Systems", Management Science, Vol. 18, No. 12, pp. 676-690.

37. Patel, M. S. [1962], "Group Screening with More than Two Stages", Technometrics, Vol. 4, No. 2, pp. 209-217.

38. Plackett, R. L. and Burman, J. P. [1946], "The Design of Optimum Multifactorial Experiments", Biometrica, Vol. 33, pp. 305-325.

39. Satterthwaite, F. D. [1959], "Random Balance Experimentation", Technometrics, Vol. 1, No. 2, pp. 111.

40. Schruben, L. W. and B. H. Margolin [1978], "Pseudorandom Number Assignment in Statistically Designed Simulation and Distribution Sampling Experiments", Journal of the American Statistical Association, Vol. 73, No. 363, pp. 504-520.

41. Watson, G. S. [1961], "A Study of the Group Screening Method", Technometrics, Vol. 3, No. 3, pp. 371-388.

42. Webb, S. R. [1965], Design, Testing and Estimation in Complex Experimentation, Office of Aerospace Research, ARL 65-116.

43. Webb, S. R. [1968], "Saturated Sequential Factorial Designs", Technometrics, Vol. 10, No. 3, pp. 535-550.

44. Webb, S. R. [1971], "Small Incomplete Factorial Experimental Designs for Two-Level and Three-Level Factors", Technometrics, Vol. 13, No. 2, pp. 243-256.

45. Williams, D. R. and D. L. Weeks [1974], "A Technique for Designating and Augmenting Computer Simulation Experiments", Management Science, Vol. 20, No. 10.